

GenCore version 5.1.7  
Copyright (c) 1993 - 2006 Bioceleration Ltd.

OM protein - protein search, using sw model

Run on: February 4, 2006, 12:09:23 ; Search time 199 Seconds  
(without alignments)  
397.428 Million cell updates/sec

Title: US-09-529-206E-4

Perfect score: 959

Sequence: 1 MQAEGRTGTGSDADGPG.....WITQCFLPVFLAQPSPGQR 180

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 2443163 seqs, 439378781 residues

Total number of hits satisfying chosen parameters: 2443163

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

- A\_Geneseq\_21:\*\*  
1: Geneseqp1980s:\*\*  
2: Geneseqp1990s:\*\*  
3: Geneseqp2000s:\*\*  
4: Geneseqp2001s:\*\*  
5: Geneseqp2002s:\*\*  
6: Geneseqp2003as:\*\*  
7: Geneseqp2003bs:\*\*  
8: Geneseqp2004s:\*\*  
9: Geneseqp2005s:\*\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

# SUMMARIES

Result No.	Score	Match	Length	DB	ID	Description
1	959	100.0	180	2	AAW62584	Cancer as
2	959	100.0	180	2	AAW69665	Human NY-
3	959	100.0	180	2	AAY05965	Human can
4	959	100.0	180	3	AAY52430	Human tum
5	959	100.0	180	3	AAY70862	Human oes
6	959	100.0	180	3	AAW69946	Human NY-
7	959	100.0	180	4	AAW67164	Amino aci
8	959	100.0	180	4	AAU01535	Human NY-
9	959	100.0	180	4	AAW07714	Human NY
10	959	100.0	180	5	AAU84818	Human NYN
11	959	100.0	180	5	AAU11543	Human tum
12	959	100.0	180	5	AAU11543	Human tum
13	959	100.0	180	6	ABR58672	Human can
14	959	100.0	180	6	ABR48210	Human bla
15	959	100.0	180	6	ABU56508	Lung canc
16	959	100.0	180	6	ABU56594	Lung canc
17	959	100.0	180	6	ABP74198	Human NY-
18	959	100.0	180	6	ABR83438	Human NY-
19	959	100.0	180	7	ADC09576	NY-ESO-1
20	959	100.0	180	7	ADD35568	Human NY-
21	959	100.0	180	7	ADD25510	Binding d
22	959	100.0	180	7	ADN39068	Cancer/an
23	959	100.0	180	8	ADJ54139	Human NY-
24	959	100.0	180	8	ADM72815	Human NY-

25	959	100.0	180	8	ADQ18451	Human sof
26	959	100.0	180	8	ADQ10446	Autoimmun
27	959	100.0	180	8	ADS80926	Tumour as
28	959	100.0	180	9	ADW44353	Human aut
29	959	100.0	180	9	ADY85096	Tumor ant
30	959	100.0	180	9	AEA35651	Human NY-
31	959	100.0	397	4	AAE13122	NY-ESO-1/
32	954	99.5	240	9	ADM99402	NY-ESO-1/
33	954	99.5	240	9	AE880047	Human NY-
34	953	98.4	180	6	AEU64816	Human NY-
35	953	99.4	180	9	ADZ28913	NY-ESO-1
36	953	99.4	180	9	ADZ42374	Immunogen
37	950	99.1	180	8	ADM73418	CAG-3 pro
38	950	99.1	180	8	ADM73417	Human NY-
39	934.5	97.4	179	8	ADK68648	Epitope 1
40	918	95.7	180	7	ADD35564	Human NY-
41	821	85.6	180	6	ABP74199	Human LAG
42	821	85.6	180	7	ADC09577	LAGR-la p
43	821	85.6	180	8	ADM72816	Human LAG
44	814	84.9	180	2	AAW69664	Human LAG
45	814	84.9	180	3	AAY70860	Human LAG

## ALIGNMENTS

RESULT 1  
AAW62584  
ID AAW62584 standard; protein; 180 AA.

XX AC AAW62584;

DT 17-SEP-1998 (first entry)

DE Cancer associated antigen NY-ESO-1.

XX KW Cancer associated antigen; NY-ESO-1; regression; progression; onset;

XX KW cancer; treatment; diagnosis.

OS Homo sapiens.

XX FH Key Location/Qualifiers

FT Misc-difference 7 /note= "potential myristoylation site"

FT Misc-difference 9 /note= "potential myristoylation site"

FT Misc-difference 11 /note= "potential phosphorylation site"

FT Misc-difference 98 /note= "potential phosphorylation site"

FT Misc-difference 134 /note= "potential phosphorylation site"

FT Misc-difference 138 /note= "potential phosphorylation site"

XX WO9814464-A1.

PD 09-APR-1998.

PF 15-SEP-1997; 97WO-US016335.

XX 03-OCT-1996; 96US-00725182.

XX (LUDW-) LUDWIG INST CANCER RES.

XX Chen Y, Scanlan M, Gure A, Old LJ, Jager E, Knuth A;

PI Driffoht JW;

XX WPI; 1998-286417/25.

XX N-PSDB; AAV38566.

PT New isolated cancer associated antigen - is used to develop products for

the diagnosis and treatment of cancers and for monitoring cancer therapy.



XX The present sequence represents the ORF1 protein encoded by open reading  
CC frame 1 of the human ESO-1/CAG-3 (or CAG-3) gene. CAG-3 is a new and  
CC potent tumour antigen capable of eliciting an antigen specific immune  
CC response by T cells. Cancer peptides comprising ORF1, ORF2 (see  
CC AAY05966), portions of these peptides and their variants (see AAY05965-  
CC 87), are useful as cancer vaccines that protect the recipient from  
CC development of cancer. The invention provides: vectors and host cells  
CC (also useful as vaccines); a method of diagnosis of cancer or precancer;  
CC a transgenic animal; antisense oligonucleotides that inhibit expression  
CC of the cancer peptide or tumour antigen; antibodies reacting with the CAG  
CC -3 cancer peptide, useful in diagnostic and detection assays; and methods  
CC for preventing or inhibiting cancer by administering a cancer peptide,  
CC with or without an HLA molecule. The cancer peptides form part of, or are  
CC derived from, cancers such as primary or metastatic melanoma, thymoma,  
CC lymphoma, sarcoma, lung cancer, liver cancer, leukaemia, uterine cancer,  
CC cervical cancer, bladder cancer, kidney cancer and adenocarcinomas such  
CC as breast, prostate, ovarian, pancreatic and thyroid cancers. Melanoma is  
CC treated by inducing cancer-specific T cells in vitro for subsequent  
CC return to a patient  
XX  
SQ Sequence 180 AA;  
  
Query Match 100.0%; Score 959; DB 2; Length 180;  
Best Local Similarity 100.0%; Pred. No. 5.9e-79;  
Matches 180; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
  
QY 1 MQAEGRTGSGTDADGPGGPGTDPGCGNAGGPGAGATGCGPRGAGARASGPGGA 60  
Db |||||||  
1 MQAEGRTGSGTDADGPGGPGTDPGCGNAGGPGAGATGCGPRGAGARASGPGGA 60  
  
QY 61 PRGPHGGAASGLNGCCRCGARGPESRLLEFYLAMPATPWEALARRSLAQDAPPLVPFG 120  
Db |||||||  
61 PRGPHGGAASGLNGCCRCGARGPESRLLEFYLAMPATPWEALARRSLAQDAPPLVPFG 120  
  
QY 121 VLKKEVTSGNLTIRLTAAHRQLQSLSSCLQLSLLMWITQCPLPVLQPPSGQR 180  
Db |||||||  
121 VLKKEVTSGNLTIRLTAAHRQLQSLSSCLQLSLLMWITQCPLPVLQPPSGQR 180  
  
RESULT 4  
AAY52430  
ID AAY52430 standard; protein; 180 AA.  
XX  
AC AAY52430;  
XX  
DT 21-OCT-2004 (revised)  
DT 15-FEB-2000 (first entry)  
XX  
DE Human tumour antigen NY-ESO-1.  
XX  
KW Cancer; tumour; antigen; MHC; major histocompatibility complex; T-cell;  
KW cytotoxic; helper; stimulation; proliferation; treatment; diagnosis;  
KW prevention; melanoma; breast cancer; ovarian cancer; prostate cancer;  
KW hepatoma; thyroid cancer; bladder cancer; lung cancer; lymphoma.  
XX  
OS Homo sapiens.  
OS Unidentified.  
XX  
XX Key Location/Qualifiers  
FH Peptide 44..53  
FT /note= "Peptide presented by MHC Class I HLA-B7"  
FT Peptide 60..69  
FT /note= "Peptide presented by MHC Class I HLA-B7"  
FT Peptide 60..68  
FT /note= "Peptide presented by MHC Class I HLA-B7"  
FT Peptide 63..72  
FT /note= "Peptide presented by MHC Class I HLA-B7"  
FT Peptide 79..88  
FT /note= "Peptide presented by MHC Class I HLA-B7"  
FT Peptide 79..87  
FT /note= "Peptide presented by MHC Class I HLA-B7, HLA-B8  
FT and HLA-B35"

FT Peptide 82..91  
FT /note= "Peptide presented by MHC Class I HLA-A1"  
FT Peptide 82..90  
FT /note= "Peptide presented by MHC Class I HLA-A1"  
FT Peptide 83..91  
FT /note= "Peptide presented by MHC Class I HLA-B44"  
FT Peptide 84..92  
FT /note= "Peptide presented by MHC Class I HLA-B7, HLA-B8  
FT and HLA-B35"  
FT Peptide 87..96  
FT /note= "Peptide presented by MHC Class I HLA-A1"  
FT Peptide 88..96  
FT /note= "Peptide presented by MHC Class I HLA-B44"  
FT Peptide 96..104  
FT /note= "Peptide presented by MHC Class I HLA-B7"  
FT Peptide 100..108  
FT /note= "Peptide presented by MHC Class I HLA-B44"  
FT Peptide 102..110  
FT /note= "Peptide presented by MHC Class I HLA-B44"  
FT Peptide 107..116  
FT /note= "Peptide presented by MHC Class I HLA-A24"  
FT Peptide 110..118  
FT /note= "Peptide presented by MHC Class I HLA-B52"  
FT Peptide 113..122  
FT /note= "Peptide presented by MHC Class I HLA-B7 and HLA-  
FT B52"  
FT Peptide 113..121  
FT /note= "Peptide presented by MHC Class I HLA-B7"  
FT Peptide 115..124  
FT /note= "Peptide presented by MHC Class I HLA-A3"  
FT Peptide 118..126  
FT /note= "Peptide presented by MHC Class I HLA-B35"  
FT Peptide 124..133  
FT /note= "Peptide presented by MHC Class I HLA-B52"  
FT Peptide 125..133  
FT /note= "Peptide presented by MHC Class I HLA-A24"  
FT Peptide 138..147  
FT /note= "Peptide presented by MHC Class I HLA-B8"  
FT Peptide 139..147  
FT /note= "Peptide presented by MHC Class I HLA-B7"  
FT Peptide 145..153  
FT /note= "Peptide presented by MHC Class I HLA-A24 and HLA-  
FT B52"  
FT Peptide 153..162  
FT /note= "Peptide presented by MHC Class I HLA-B52"  
FT Peptide 154..163  
FT /note= "Peptide presented by MHC Class I HLA-B52"  
FT Peptide 154..162  
FT /note= "Peptide presented by MHC Class I HLA-B52"  
FT Peptide 156..167  
FT /note= "Peptide (AAY52434) presented by MHC Class I HLA-  
FT A2"  
FT Peptide 158..166  
FT /note= "Peptide presented by MHC Class I HLA-A3"  
FT Peptide 159..167  
FT /note= "Peptide presented by MHC Class I HLA-A3"  
FT Peptide 162..170  
FT /note= "Peptide presented by MHC Class I HLA-B52"  
FT  
XX WO953938-A1.  
PN  
XX  
PD 28-OCT-1999.  
XX  
XX 24-MAR-1999; 99WO-US006875.  
XX  
PR 17-APR-1998; 98US-00062422.  
PR 02-OCT-1998; 98US-00165546.  
XX  
XX (LUDW-) LUDWIG INST CANCER RES.  
XX  
XX Stockert E, Jager E, Chen Y, Scanlan M, Alexander K, Old LJ;  
PI Gure A, Ritter G;  
XX

```
DR WPI; 2000-038483/03.
XX N-PSDB; AAZ38380.
XX Novel peptides which bind to MHC class I and MHC class II molecules,
XX useful for therapeutic and diagnostic purposes.
XX Claim 30; Fig 3; 49pp; English.
XX This sequence represents a human tumour antigen, NY-ESO-1, the cDNA
XX encoding which was isolated from an oesophagus squamous cell cancer cDNA
XX library. Tissue localisation studies revealed it to be expressed at high
XX levels in normal ovary and testis but not in normal colon, kidney, liver,
XX brain, oesophagus and skin. It was expressed in certain tumours and
XX tumour cell lines with some degree of frequency - these included melanoma
XX specimens and cell lines, and breast and bladder cancer specimens, with
XX expression in other tumour types being sporadic. Peptides derived from NY
XX -ESO-1 are bound by both MHC (major histocompatibility complex) Class I
XX and Class II molecules for presentation to T-cells. Peptides AAY52431-
XX Y52434 bind to Class I HLA-A2 molecules, thereby stimulating
XX proliferation of cytotoxic T-cells, while peptides AAY52435-Y52440 bind
XX to Class II HLA-DR53 molecules, stimulating helper T-cell proliferation.
XX The peptides derived from NY-ESO-1 may be used in methods and
XX compositions used for the treatment, diagnosis and prevention of cancers
XX (such as melanoma, breast cancer, prostate cancer, lung cancer, hepatoma,
XX ovarian cancer, thyroid cancer, bladder cancer, or lymphoma) and to
XX stimulate the proliferation of T cells
XX Revised record issued on 21-OCT-2004 : Correction to feature table key
XX Sequence 180 AA;
XX Query Match 100.0%; Score 959; DB 3; Length 180;
XX Best Local Similarity 100.0%; Pred. No. 5.9e-79;
XX Matches 180; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1 MQAEGRTGGSTGDADGGPGGIPDGGNAGGPGGAGATGGPRGAGAAARASGPGGGA 60
Db 1 MQAEGRTGGSTGDADGGPGGIPDGGNAGGPGGAGATGGPRGAGAAARASGPGGGA 60
QY 61 PRGPHGGAASGLNGCCRCGARGPESRLLEFYLLAMPFATPMEALARRSLAQDAPPLVPVG 120
Db 61 PRGPHGGAASGLNGCCRCGARGPESRLLEFYLLAMPFATPMEALARRSLAQDAPPLVPVG 120
QY 121 VLLKEFTVSGNLTIRLTAADHRQLQLSISSCLQLSLLMWITQCFLPVFLAQPSPGQRR 180
Db 121 VLLKEFTVSGNLTIRLTAADHRQLQLSISSCLQLSLLMWITQCFLPVFLAQPSPGQRR 180
RESULT 5
AAY70862
ID AAY70862 standard; protein; 180 AA.
XX AC AAY70862;
XX DT 31-JUL-2000 (first entry)
XX DE Human tumour antigen, NY-ESO-1 protein.
XX KW NY-ESO-1; CAMEL; CTL-recognised Antigen on MELANOMA; human; cancer; CTL;
XX melanoma; immunotherapy; immune response.
XX OS Homo sapiens.
XX FT Modified-site 7
XX FT Modified-site 9
XX FT Modified-site 11
XX FT Modified-site 98
XX FT Modified-site 134
XX FT Modified-site 138
XX (BOEH ) BOEHRINGER INGELHEIM INT GMBH.

PA (UYHO-) UNIV HOSPITAL LEIDEN.
XX Schrier PI, Aarnoudse CA, Heider K, Klade C;
XX WPI; 2000-339685/29.
XX N-PSDB; AAD00152.
XX Tumor-associated antigen useful for cancer immunotherapy is encoded by
XX the open reading frame of LAGE-1 (a tumor-specific antigen) cDNA.
XX Example 3; Page 62-63; 73pp; English.
XX The present sequence is the human NY-ESO-1 protein, a tumour antigen,
XX identified by screening an esophagus carcinoma cDNA library. This protein
XX is derived from open reading frame (ORF)-1 that contain epitopes of
XX tumour specific T-cells. NY-ESO-1 is expressed in different tumour types,
XX but not in healthy tissues except in testis. It also shows homology with
XX the CAMEL (Cytotoxic T lymphocytes (CTL)-recognised Antigen on MELANOMA)
XX protein, a tumour-associated antigen. The tumour-associated antigen
XX This sequence has anticancer activity. CAMEL tumour antigen and
XX immunogenic peptides derived from it are useful for cancer immunotherapy.
XX They have the potential to induce an immune response, by eliciting a CTL
XX response. The DNA molecule is used for the construction of recombinant or
XX fusion proteins
XX Sequence 180 AA;
XX Query Match 100.0%; Score 959; DB 3; Length 180;
XX Best Local Similarity 100.0%; Pred. No. 5.9e-79;
XX Matches 180; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1 MQAEGRTGGSTGDADGGPGGIPDGGNAGGPGGAGATGGPRGAGAAARASGPGGGA 60
Db 1 MQAEGRTGGSTGDADGGPGGIPDGGNAGGPGGAGATGGPRGAGAAARASGPGGGA 60
QY 61 PRGPHGGAASGLNGCCRCGARGPESRLLEFYLLAMPFATPMEALARRSLAQDAPPLVPVG 120
Db 61 PRGPHGGAASGLNGCCRCGARGPESRLLEFYLLAMPFATPMEALARRSLAQDAPPLVPVG 120
QY 121 VLLKEFTVSGNLTIRLTAADHRQLQLSISSCLQLSLLMWITQCFLPVFLAQPSPGQRR 180
Db 121 VLLKEFTVSGNLTIRLTAADHRQLQLSISSCLQLSLLMWITQCFLPVFLAQPSPGQRR 180
RESULT 6
AAB03154
ID AAB03154 standard; protein; 180 AA.
XX AC AAB03154;
XX DT 23-OCT-2000 (first entry)
XX DE Human oesophageal cancer-associated antigen NY-ESO-1.
XX KW Oesophageal cancer associated antigen; NY-ESO-1; human; immunogen;
XX oesophageal carcinoma; melanoma; ovary; testis; transmembrane domain;
XX antibody; diagnostic marker; drug delivery target.
XX OS Homo sapiens.
XX FT Modified-site 7
XX FT Modified-site 9
XX FT Modified-site 11
XX FT Modified-site 98
XX FT Modified-site 134
XX FT Modified-site 138
XX Key Location/Qualifiers
XX /note= "Potential N-myristoylation site"
XX /note= "Potential N-myristoylation site"
XX /note= "Potential O-phosphorylation site"
XX /note= "Potential O-phosphorylation site"
XX /note= "Potential O-phosphorylation site"
XX /note= "Potential O-phosphorylation site"
```





```
AAE07714 standard; protein; 180 AA.
AAE07714;
06-NOV-2001 (first entry)
Human NY ESO-1 protein.
Human; major histocompatibility complex; MHC; vaccine; metastasis;
class II restricted T cell epitope; MHC-II epitope; cancer antigen;
NY ESO-1 protein; CD4+ T lymphocyte; human leucocyte antigen; HLA;
tumour-specific humoral-mediated immunity; cancer; cytostatic;
immunotherapy.
Homo sapiens.
Key Location/Qualifiers
Misc-difference 45..47 /note= "Encoded by CCCGGGGC"
WO200155393-A2.
02-AUG-2001.
26-JAN-2001; 2001WO-US002765.
28-JAN-2000; 2000US-0179004P.
29-SEP-2000; 2000US-0237107P.
(USSH ) US DEPT HEALTH & HUMAN SERVICES.
Wang R, Rosenberg SA, Zeng G;
WPI: 2001-496851/54.
N-PSDB; AAD14179, AAD14180.
New NY-ESO cancer peptide or MHC class II restricted T cell epitopes,
useful as immunogen and vaccine for inhibiting cancer in a mammal or as
protection from metastasis.
Example 1; Fig 1; 134pp; English.
The invention relates to the identification and isolation of major
histocompatibility (MHC) class II restricted T cell epitope (MHC-II
epitope) derived from the cancer antigen, NY ESO-1. The MHC-II epitopes
from NY ESO-1 are recognised by CD4+ T lymphocytes in an human leucocyte
antigen (HLA) class II restricted manner, in particular HLA-DR or HLA-DP
restricted. The products of the gene are promising candidates for
immunotherapeutic strategies for the prevention, treatment and diagnosis
of patients with cancer. The cancer epitopes are useful as immunogen and
vaccine to inhibit or to prevent cancer in a mammal by eliciting CD4+ T
lymphocytes resulting in protection of the recipient from development of
cancer and protection from metastasis, or by inhibiting the growth of
cells expressing the NY-ESO-1 gene product. The cancer peptides are also
useful as diagnostic agent to detect the presence of cancer, to enhance
the generation of antibody and/or CD8+ T cell responses against any given
target antigen and/or hapten and to induce tumour-specific humoral-
mediated immunity against cancer. The present sequence is human NY ESO-1
protein
Sequence 180 AA;
Query Match 100.0%; Score 959; DB 4; Length 180;
Best Local Similarity 100.0%; Pred. No. 5.9e-79;
Matches 180; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1 MQAEGRTGGTGDADPGGPGIIPDGGNAGGPGGAGATGCGPRGAGAAASGPGGA 60
DB 1 MQAEGRTGGTGDADPGGPGIIPDGGNAGGPGGAGATGCGPRGAGAAASGPGGA 60
QY 61 PRGPHGASGLNGCCRCGARGPESLLEFLYAMPATPWEALRRSLAQADPLPVG 120
DB 61 PRGPHGASGLNGCCRCGARGPESLLEFLYAMPATPWEALRRSLAQADPLPVG 120
```

```
121 VLLKEFTVSGNLTITRLTAADHRQLQLSISSCLQQLSLLMWITQCFLPVFLAQPSPGQRR 180
121 VLLKEFTVSGNLTITRLTAADHRQLQLSISSCLQQLSLLMWITQCFLPVFLAQPSPGQRR 180
RESULT 11
AAU84818
ID AAU84818 standard; protein; 180 AA.
XX
AC AAU84818;
XX
DT 08-MAY-2002 (first entry)
XX
DE Human NYNSO1a consensus sequence.
XX
KW Savine; vaccine; cancer; viral infection; HIV; hepatitis C virus;
KW viral infection; human immunodeficiency virus; melanoma;
KW bacterial infection; Salmonella; Legionella; parasitic infection;
KW Trypanosoma; Toxoplasma; Giardia.
XX
OS Homo sapiens.
XX
PN WO200190197-A1.
XX
PD 29-NOV-2001.
XX
PF 25-MAY-2001; 2001WO-AU000622.
XX
PR 26-MAY-2000; 2000AU-00007761.
XX
PA (AUSU ) UNIV AUSTRALIAN NAT.
XX
PI Thomson SA, Ramshaw IA;
XX
WPI: 2002-147575/19.
XX
New synthetic polypeptides having several different segments of at least
one parent polypeptide linked together differently compared to the
linkage in the parent polypeptide, for inducing immune response against a
pathogen or cancer.
Example 3; Fig 27; 364pp; English.
The invention relates to a new synthetic polypeptide (I) comprising
several different segments of at least one parent polypeptide linked
together in a different relationship relative to their linkage in the
parent polypeptide to impede, abrogate or otherwise alter at least one
function associated with the parent polypeptide and for inducing an
immune response against a pathogen or cancer. Also included are a
synthetic polynucleotide encoding and a computer system for designing the
synthetic polypeptides. The synthetic polypeptides and polynucleotides
are referred to as a Savine. The synthetic polypeptide is useful for
modulating immune responses preferably directed against a pathogen or a
cancer, (e.g., cancers of the lung, breast, ovary, cervix, colon, head
and neck, pancreas, prostate, stomach, bladder, kidney, bone liver,
oesophagus, brain, testicle, uterus), as potentiating agents.
Compositions comprising the polypeptide may be used in the treatment or
prophylaxis against viral (such as infections caused by HIV (human
immunodeficiency virus), hepatitis, influenza, Japanese encephalitis
virus, Epstein-Barr virus and respiratory syncytial virus), bacterial
(e.g., infections caused by Neisseria, Meningococcal, Haemophilus,
Salmonella, Streptococcal, Legionella and Mycobacterium or parasitic
(e.g., infections caused by Plasmodium, Schistosoma, Leishmania,
Trypanosoma, Toxoplasma and Giardia) infections. The present sequence is
a consensus sequence for a parent protein used to design a savine of the
invention
Sequence 180 AA;
Query Match 100.0%; Score 959; DB 5; Length 180;
Best Local Similarity 100.0%; Pred. No. 5.9e-79;
Matches 180; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
```

QY	1	MQAEGRTGGTGDADGPGGPGI	PDGPGNAGGPGCEAGATGGRGPRGAGAAASGPGGGA	60
Db	1	MQAEGRTGGTGDADGPGGPGI	PDGPGNAGGPGCEAGATGGRGPRGAGAAASGPGGGA	60
QY	61	PRGPHGGAASGLNGCCRCGARGPSRLLEFYLAMPFATPMEAEIARRSLAODAPPLVPVG	120	
Db	61	PRGPHGGAASGLNGCCRCGARGPSRLLEFYLAMPFATPMEAEIARRSLAODAPPLVPVG	120	
QY	121	VLLKPTVSGNLTIRLTAADHRQLQLSIS	CLQLSISCLQLSLLMWITCFLPVLAPPPGQRR	180
Db	121	VLLKPTVSGNLTIRLTAADHRQLQLSIS	CLQLSLLMWITCFLPVLAPPPGQRR	180
RESULT 12				
AAU11543				
II	AAU11543	standard; protein; 180 AA.		
XX	AAU11543;			
AC	XX			
DT	12-MAR-2002	(first entry)		
XX	XX	Human tumour associated antigen NY-ESO.		
DE	XX			
KW	KW	Human; tumour associated antigen; NY-ESO; human leukocyte antigen; major histocompatibility complex; MHC; HLA-A2.2; vaccine; cancer; HIV; human immunodeficiency virus infection; cytostatic; virucide; housekeeping epitope; adoptive immunotherapy; neoplastic disease; viral disease; hepatitis virus; papilloma virus; tumour; leukaemia; lymphoma; breast cancer; prostate cancer; lung cancer; parasitic infection; Chlamydia; Trypanosoma; Toxoplasma.		
XX	OS	Homo sapiens.		
XX	XX			
PN	WO200182963-A2.			
XX	XX			
PD	08-NOV-2001.			
XX	XX			
PF	27-APR-2001; 2001WO-US013806.			
XX	XX			
PR	28-APR-2000; 2000US-00560465.			
PR	28-APR-2000; 2000US-00561074.			
PR	28-APR-2000; 2000US-00561571.			
PR	28-APR-2000; 2000US-00561572.			
XX	XX	(CTLI-) CTL IMMUNOTHERAPIES CORP.		
PA	XX			
XX	XX			
P:	Simard JUL, Diamond DC, Lei X;			
XX	XX			
DR	WPI; 2002-0666492/09.			
XX	XX			
P:	Novel vaccine useful for treating neoplastic and viral diseases, comprises a first housekeeping epitope derived from a first antigen associated with a first target cell.			
XX	XX			
PS	Example 23; Fig 17; i31pp; English.			
XX	XX			
CC	CC	The invention relates to a vaccine comprising a first housekeeping epitope derived from a first antigen associated with a first target cell. Also included are an isolated T cell expressing a T cell receptor specific for a major histocompatibility complex (MHC)-peptide complex comprising a first housekeeping epitope which is derived from a first antigen associated with a first target cell, selecting an epitope (or peptide sequence) from a population of peptide fragments of an antigen associated with a target in a host, where the fragments have a known or predicted affinity for a MHC receptor peptide binding cleft of the host, where the epitope selected corresponds to a product of proteolytic cleavage of the antigen in a cell of the host and a nucleic acid construct comprising a first coding region, where the first coding region comprises a first sequence encoding at least a first polypeptide, where the first polypeptide comprises a first housekeeping epitope derived from a first antigen associated with a first target cell; The epitopes, peptides, vaccines and nucleic acids are useful in the manufacture of a		

CC	medicament for use in adoptive immunotherapy and for prevention and
CC	treatment of neoplastic and viral diseases (e.g. human immunodeficiency
CC	virus, HIV, infection, hepatitis virus and papilloma virus), cancers
CC	(e.g. tumours), leukaemia, lymphoma, breast cancer, prostate cancer and
CC	lung cancer), infection of cells by intracellular parasites (e.g.
CC	Chlamydia, Trypanosoma and Toxoplasma) and many other examples given in
CC	the specification. The invention permits the vaccine designer to ignore
CC	peptides that, despite predicted high binding affinity for MHC, will
CC	never be useful because they cannot be presented by target cells. The
CC	invention provides a major advance in vaccine design, one that combines
CC	the power of antigen sequence analysis with the fundamental realities of
CC	immunology. The invention allows for the simple and effective selection
CC	of meaningful epitopes for creation of MHC class I or Class II vaccines
CC	using any polypeptide sequence corresponding to a desired target. The
CC	present sequence is an HLA-A2.1 (human leukocyte antigen) presenting
CC	target cell protein from which epitopes of the invention may be derived,
CC	NY-ESO (a tumour associated antigen)
XX	
SQ	Sequence 180 AA;
	Query Match            100.0%;    Score 959;    DB 5;    Length 180;
	Best Local Similarity 100.0%;    Pred. No. 5.9e-79;
	Matches 180; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY	1 MQAEGRTGSGTGDADPGCGGPIPDGPGGNAGGPGAGATGGRGPRGAGAARASGGGGA 60
DB	
DB	1 MQAEGRTGSGTGDADPGCGGPIPDGPGGNAGGPGAGATGGRGPRGAGAARASGGGGA 60
QY	61 PRPHGGAASGLNGCCRCGARGPESRLLEPYLAMPFATPWEAEIARRSLAQDAPPLVPVG 120
DB	
DB	61 PRPHGGAASGLNGCCRCGARGPESRLLEPYLAMPFATPWEAEIARRSLAQDAPPLVPVG 120
QY	121 VLKFTVTSNIIITIRLTAAHROLQLSISSCLQQLSLLMWTTCFLPVFLAQPSPGQR 180
DB	
DB	121 VLKFTVTSNIIITIRLTAAHROLQLSISSCLQQLSLLMWTTCFLPVFLAQPSPGQR 180
RESULT 13	
ABR58672	
ID	ABR58672 standard; protein; 180 AA.
AC	ABR58672;
DT	09-JUL-2003 (first entry)
DE	Human cancer related protein SEQ ID NO:329.
KW	Human; cancer; diagnosis; screening; modulator; leukaemia; ischaemia;
KW	heart disease; atherosclerosis; endometriosis.
OS	Homo sapiens.
PN	WO2003025138-A2.
FD	27-MAR-2003.
PP	17-SEP-2002; 2002WO-US029560.
FR	17-SEP-2001; 2001US-0323469P.
PR	20-SEP-2001; 2001US-0323887P.
PR	13-NOV-2001; 2001US-0350666P.
PR	08-FEB-2002; 2002US-0355145P.
PR	18-FEB-2002; 2002US-0355257P.
PR	12-APR-2002; 2002US-0372246P.
PA	(EOSB-) EOS BIOTECHNOLOGY INC.
XX	
FI	Afar D, Aziz N, Gish KC, Hevezi PA, Mack DH, Wilson KE;
FI	Zlotnik A;
DR	WPI; 2003-354600/33.
DR	N-PSDB; ACCT2823.
XX	

PT New genes that are up-regulated or down-regulated in cancers, useful as  
PT markers for diagnosing e.g. cancer, ischemia or heart diseases, or as  
PT therapeutic targets for screening drugs for treating these diseases.  
XX  
PS Claim 12; Page 757-758; 767pp; English.

XX The present invention describes an isolated nucleic acid molecule, which  
CC comprises the sequence of any of the genes that are up-regulated or down-  
CC regulated in specific cancers (e.g. about 1031 genes up-regulated in  
CC acute lymphocytic leukemia). ACC72641 to ACC72860 represent cancer  
CC related gene nucleotide sequences which encode the proteins given in  
CC ABR58521 to ABR58709. Also described: (1) determining the presence or  
CC absence of a pathological cell in a patient; (2) an expression vector  
CC comprising a nucleic acid molecule described above; (3) a host cell  
CC comprising the vector; (4) an isolated polypeptide, which is encoded by  
CC the nucleic acid; (5) an antibody that specifically binds the polypeptide  
CC of (4); (6) specifically targeting a compound to a pathological cell in a  
CC patient by administering to the patient the antibody above; and (7) a  
CC drug screening assay. The nucleic acid is useful as diagnostic markers or  
CC therapeutic targets. In particular, the nucleic acid is useful for  
CC diagnosing a pathology, e.g. cancer (e.g. cancer of the bone marrow,  
CC bladder, brain, breast, cervix, colon/rectum, kidney, lung, ovary,  
CC pancreas, prostate, skin and uterus), wounds, ischaemia, heart diseases,  
CC atherosclerosis and endometriosis. The nucleic acid is also useful in  
CC drug screening, particularly for identifying agents for treating these  
CC pathologies  
XX  
SQ Sequence 180 AA;

Query Match 100.0%; Score 959; DB 6; Length 180;  
Best Local Similarity 100.0%; Pred. No. 5.9e-79;  
Matches 180; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 MQAEGRTGCGTGADGPGGPGIPDGGNAGGPGGAGATGCGPRGAGAAASGPGGA 60  
DB 1 MQAEGRTGCGTGADGPGGPGIPDGGNAGGPGGAGATGCGPRGAGAAASGPGGA 60  
QY 61 PRGPHGGAASGLNGCCCGARGPESRLLEFYLLAMPATPMEAEIARRSLAQDAPPLVPVG 120  
DB 61 PRGPHGGAASGLNGCCCGARGPESRLLEFYLLAMPATPMEAEIARRSLAQDAPPLVPVG 120  
QY 121 VLLKEFTVSGNLTIRLTAAADHRQLQSLSSCQLQLSLLMWITQCFLPVFLAQPSPGQRR 180  
DB 121 VLLKEFTVSGNLTIRLTAAADHRQLQSLSSCQLQLSLLMWITQCFLPVFLAQPSPGQRR 180

RESULT 14  
ABR48210  
ID ABR48210 standard; protein; 180 AA.  
AC ABR48210;

DT 12-JUN-2003 (first entry)

DE Human bladder cancer associated protein sequence SEQ ID NO:139.

XX Human; bladder cancer; cytostatic; gene therapy; vaccine.

XX Homo sapiens.

OS WO2003003906-A2.

XX 16-JAN-2003.

XX 03-JUL-2002; 2002WO-US021338.

XX 03-JUL-2001; 2001US-0302814P.

PR 03-AUG-2001; 2001US-0310099P.

PR 08-NOV-2001; 2001US-0343705P.

PR 13-NOV-2001; 2001US-0350666P.

PR 12-APR-2002; 2002US-0372246P.

XX (EOSB-) EOS BIOTECHNOLOGY INC.

XX Mack DH, Aziz N;  
XX WPI; 2003-201532/19.  
DR N-PSDB; ACC51024.

XX Detecting a bladder cancer-associated transcript in a cell from a  
PT patient, comprises contacting a biological sample from the patient with a  
PT bladder cancer-associated polynucleotide or antibody.

PS Claim 10; Page 278; 307pp; English.

XX The present invention describes a method for detecting a bladder cancer-  
CC associated transcript in a cell from a patient. The method comprises  
CC contacting a biological sample from the patient with a polynucleotide  
CC that selectively hybridises to a sequence that is 80 % identical to a  
CC table of sequences (see ACC50951 to ACC51059). ACC50951 to ACC51059  
CC encode the human bladder cancer-associated proteins given in ABR48146 to  
CC ABR48242). Bladder cancer-associated sequences from the present invention  
CC have cytostatic activities, and can be used in antisense gene therapy and  
CC in vaccine production. The method can be used for detecting a bladder  
CC cancer-associated transcript in a cell from a patient. The method is  
CC useful in diagnosing or treating bladder cancer and in screening for  
CC compounds that modulate bladder cancer, such as hormones or antibodies.  
CC The nucleic acid molecules from the present invention may be used in  
CC various screening and diagnostic methods, and for gene therapy, vaccine  
CC and/or antisense/inhibition applications  
XX  
SQ Sequence 180 AA;

Query Match 100.0%; Score 959; DB 6; Length 180;  
Best Local Similarity 100.0%; Pred. No. 5.9e-79;  
Matches 180; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 MQAEGRTGCGTGADGPGGPGIPDGGNAGGPGGAGATGCGPRGAGAAASGPGGA 60  
DB 1 MQAEGRTGCGTGADGPGGPGIPDGGNAGGPGGAGATGCGPRGAGAAASGPGGA 60  
QY 61 PRGPHGGAASGLNGCCCGARGPESRLLEFYLLAMPATPMEAEIARRSLAQDAPPLVPVG 120  
DB 61 PRGPHGGAASGLNGCCCGARGPESRLLEFYLLAMPATPMEAEIARRSLAQDAPPLVPVG 120  
QY 121 VLLKEFTVSGNLTIRLTAAADHRQLQSLSSCQLQLSLLMWITQCFLPVFLAQPSPGQRR 180  
DB 121 VLLKEFTVSGNLTIRLTAAADHRQLQSLSSCQLQLSLLMWITQCFLPVFLAQPSPGQRR 180

RESULT 15  
ABU56508  
ID ABU56508 standard; protein; 180 AA.

AC ABU56508;

DT 02-APR-2003 (first entry)

DE Lung cancer-associated polypeptide #101.

XX Lung cancer-associated polypeptide; cytostatic; emphysema;  
KW antinflammatory; antiasthmatic; non-small cell lung cancer; atelectasis;  
KW small cell lung cancer; benign lesion; precancerous lesion; bronchitis;  
KW chronic obstructive pulmonary disease; hypersensitivity pneumonitis;  
KW interstitial pulmonary fibrosis; fibrosis; asthma; bronchiectasis.

OS Unidentified.

XX WO200286443-A2.

XX 31-OCT-2002.

XX 18-APR-2002; 2002WO-US012476.

PR 18-APR-2001; 2001US-0284770P.

PR 10-MAY-2001; 2001US-0290492P.



GenCore version 5.1.7  
Copyright (c) 1993 - 2006 Bioceleration Ltd.

OM protein - protein search, using sw model

Run on: February 4, 2006, 12:13:27 ; Search time 42 Seconds  
(without alignments)  
412.358 Million cell updates/sec

Title: US-09-529-206E-4

Perfect score: 959

Sequence: 1 MQAEGRTGGTGADGPGG.....WITQCFLPVFLAQPFGQRR 180

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 283416 seqs, 96216763 residues

Total number of hits satisfying chosen parameters: 283416

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

PIR 80:\*

1: pir1:\*

2: pir2:\*

3: pir3:\*

4: pir4:\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	174.5	18.2	627	T35608	polyketide hydroxy
2	155.5	16.2	145	T08435	la costa protein
3	155.5	16.2	641	Q0B331	nuclear antigen EB
4	153	16.0	1079	B70807	hypothetical glyci
5	152.5	15.9	779	CGB01S	collagen alpha 1(I
6	151.5	15.8	1464	CGH01S	collagen alpha 1(I
7	151	15.7	1489	D70807	hypothetical glyci
8	151	15.7	1901	F70806	hypothetical glyci
9	150.5	15.7	530	T32812	hypothetical prote
10	149.5	15.6	615	H70589	hypothetical glyci
11	149.5	15.6	1215	T32734	myosin-IA - Acanth
12	149	15.5	297	T27525	hypothetical prote
13	149	15.5	299	T20605	hypothetical prote
14	148.5	15.5	252	S01821	glycine-rich prote
15	148.5	15.5	1042	CGH81S	collagen alpha 1(I
16	147.5	15.4	473	I50629	collagen - chicken
17	147	15.3	296	T21070	hypothetical prote
18	147	15.3	440	F85584	probable tail comp
19	147	15.3	1391	F70806	hypothetical glyci
20	146.5	15.3	290	T24590	hypothetical prote
21	146.5	15.3	1806	1 CGH01E	collagen alpha 1(X
22	146	15.2	338	1 KNMU	glycine-rich cell
23	146	15.2	853	T70896	tractin - mediana
24	146	15.2	1880	T18531	glycine-rich prote
25	145.5	15.2	228	S19932	collagen alpha 1(X
26	145.5	15.2	482	B31795	hypothetical glyci
27	145.5	15.2	1306	T40934	hypothetical glyci
28	145	15.1	741	T70917	hypothetical glyci
29	145	15.1	1329	T70917	hypothetical glyci

30	144.5	15.1	290	2 T24586	hypothetical prote
31	144.5	15.1	301	2 B31219	collagen 2 - Caeno
32	144.5	15.1	330	2 T26004	hypothetical prote
33	144.5	15.1	438	2 D90734	probable tail fibe
34	144.5	15.1	1147	1 MWAXIB	myosin heavy chain
35	144	15.0	667	2 A70893	hypothetical glyci
36	143.5	15.0	310	2 I50696	collagen alpha 1(I
37	143.5	15.0	314	2 T32247	hypothetical prote
38	143.5	15.0	714	2 A70807	hypothetical glyci
39	143.5	15.0	886	2 I50694	collagen alpha 1(I
40	143.5	15.0	1049	1 CGH07S	collagen alpha 1(I
41	143	14.9	323	2 T19142	hypothetical prote
42	143	14.9	437	2 H90854	probable tail fibe
43	143	14.9	439	2 C90769	probable tail fibe
44	143	14.9	1660	2 A70869	hypothetical glyci
45	142.5	14.9	244	2 T49893	glycine-rich prote

ALIGNMENTS

RESULT 1

T35608

polyketide hydroxylase - Streptomyces coelicolor

C:Species: Streptomyces coelicolor

C:Date: 05-Nov-1999 #sequence\_revision 05-Nov-1999 #text\_change 09-Jul-2004

C:Accession: T35608

R:Seeger, K.J.; Harris, D.; James, K.D.; Parkhill, J.; Barrell, B.G.; Rajandream, M.A.

submitted to the EMBL Data Library, June 1999

A:Reference number: Z21584

A:Accession: T35608

A:Status: preliminary; translated from GB/EMBL/DBDJ

A:Molecule type: DNA

A:Residues: 1-627 <SEE>

A:Cross-references: UNIPROT:P42534; UNIPARC:UPI000012CFBE; EMBL:AL079356; PIDN:CAB45603.1

A:Experimental source: strain A3(2)

C:Genetics:

A:Gene: SC0EDB:SC6G9.12c

C:Superfamily: tetracycline 6-hydroxylase

Query Match 18.2%; Score 174.5; DB 2; Length 627;  
Best Local Similarity 42.7%; Pred. No. 4.1e-06;  
Matches 47; Conservative 4; Mismatches 28; Indels 31; Gaps 6;

QY 5 GRGTGTTGDADGPGGPGIPDGGNAGGPGAGATGG-RGPRGAGAAASGPGGAPRG 63  
DB 409 GRGTGG-PGGPGGGLGGPGGPGG-TGGPGGPGGPGGPRGAGAGPGGPGG---G 462

QY 64 PRGGAASGLNGCCRGCPESRLLEFYLAMPF-----ATPMEAE 103  
DB 463 PGGG-----GPGRGILNALGYRYPGAVVGADPATPVVPE 498

RESULT 2

T08435

la costa protein - fruit fly (Drosophila melanogaster)

C:Species: Drosophila melanogaster

C:Date: 11-Jun-1999 #sequence\_revision 11-Jun-1999 #text\_change 09-Jul-2004

C:Accession: T08435

R:Maloszka, R.; de Couet, H.G.; Miklos, G.L.

Proc. Natl. Acad. Sci. U.S.A. 95, 3731-3736, 1998

A:Title: Data transferability from model organisms to human beings: insights from the fur

A:Reference number: Z16415; MUID:98188272; PMID:9520435

A:Accession: T08435

A:Status: preliminary; translated from GB/EMBL/DBDJ

A:Molecule type: DNA

A:Residues: 1-145 <MAL>

A:Cross-references: UNIPROT:O61351; UNIPARC:UPI0000076890; EMBL:AF017777; NID:g3004652; I

A:Experimental source: strain Cantons

C:Genetics:

A:Gene: lcs

A:Cross-references: FlyBase:FBgn0024251

```
Query Match      16.2%; Score 155.5; DB 2; Length 145;
Best Local Similarity 51.4%; Pred. No. 3.1e-05;
Matches 37; Conservative 0; Mismatches 24; Indels 11; Gaps 4;

QY 17 GPGGPGIPDPGNAGGPGGEA-----GATGGRGPRGAGAAASGPGG-GAPRGPHGGAASG 71
    |||||
Db 25 GPGGPGGPRGRGPRGPGGPGGRGPRGPGG-----GPGGPGGPGGPGGPG 79
    |||||

QY 72 LNCRCRCRGARGP 83
    |||||
Db 80 GPG-CPGPGGP 90
    |||||

RESULT 3
Q9BE31
collagen alpha 1(I) chain - human herpesvirus 4
C:Species: human herpesvirus 4, Epstein-Barr virus
C:Date: 25-Feb-1985 #sequence revision 25-Feb-1985 #text_change 09-Jul-2004
C:Accession: C43043; S42440; A03773; S33021
R:Bankier, A.T.; Deininger, P.L.; Farrell, P.J.; Barrell, B.G.
Mol. Biol. Med. 1, 21-45, 1983
A:Title: Sequence analysis of the 17,166 bp EcoRI fragment C of B95-8 Epstein-Barr virus
A:Reference number: A93065; MUID:85035713; PMID:6092825
A:Accession: C43043
A:Molecule type: DNA
A:Residues: 1-641 <BAN>
A:Cross-references: UNIPROT:P03211; UNIPARC:UPI000000CDDDD; EMBL:V01555; NID:g59074; PIDN:
A:Experimental source: strain B95-8
R:Baer, R.; Bankier, A.T.; Biggin, M.D.; Deininger, P.L.; Farrell, P.J.; Gibson, T.J.; H
Nature 310, 207-211, 1984
A:Title: DNA sequence and expression of the B95-8 Epstein-Barr virus genome.
A:Reference number: A03794; MUID:84270667; PMID:6087149
A:Contents: annotation; protein coding region
R:Sample, J.; Hummel, M.; Braun, D.; Birkenbach, M.; Kieff, E.
Proc. Natl. Acad. Sci. U.S.A. 83, 5096-5100, 1986
A:Title: Nucleotide sequences of mRNAs encoding Epstein-Barr virus nuclear proteins: a p
A:Reference number: S42440; MUID:86259739; PMID:3460083
A:Accession: S42440
A:Status: preliminary
A:Molecule type: mRNA
A:Residues: 1-66 <SAM>
A:Cross-references: UNIPARC:UPI0000170E9C; EMBL:M13941; NID:g330399; PIDN:AAA45889.1; PI
C:Superfamily: Epstein-Barr virus nuclear antigen
C:Keywords: DNA binding; transcription regulation

Query Match      16.2%; Score 155.5; DB 1; Length 641;
Best Local Similarity 46.2%; Pred. No. 0.00012;
Matches 37; Conservative 1; Mismatches 41; Indels 1; Gaps 1;

QY 3 AEGRGTTGGTGADGPGGPIPDGPGNAGGPGGEAGATGGRGPRGAGAAASGPGGAPR 62
    |||||
Db 166 AGAGAGGGAGAGGGAGGAGGAGGAGG-AGGAGAGGAGGAGGAGGAGGAGGGA 224
    |||||

QY 63 GPHGGAASGLNGCCRCRGARG 82
    |||||
Db 225 GGAGGAGGAGGAGGAGGAGG 244
    |||||

RESULT 4
B70807
hypothetical glycine-rich protein Rv3512 - Mycobacterium tuberculosis (strain H37RV)
C:Species: Mycobacterium tuberculosis
C:Date: 17-Jul-1998 #sequence_revision 17-Jul-1998 #text_change 20-Jun-2000
C:Accession: B70807
R:Coile, S.T.; Brosch, R.; Parkhill, J.; Garnier, T.; Churcher, C.; Harris, D.; Gordon, S.
; Connor, R.T.; Davies, R.; Devlin, K.; Feltwell, T.; Gentles, S.; Hamlin, N.; Holroyd, S.
Rajandream, M.A.; Rogers, J.; Rutter, S.; Seeger, K.; Skelton, S.; Squares, S.
Nature 393, 537-544, 1998
A:Authors: Soares, R.; Suleston, J.E.; Taylor, K.; Whitehead, S.; Barrell, B.G.
A:Title: Deciphering the biology of Mycobacterium tuberculosis from the complete genome
A:Reference number: A70500; MUID:98295987; PMID:9634230
A:Accession: B70807
A:Status: preliminary; nucleic acid sequence not shown; translation not shown
```

```
A:Molecule type: DNA
A:Residues: 1-1079 <COL>
A:Cross-references: UNIPARC:UPI000000D3AEA; GB:AL022022; GB:AL123456; NID:g3261554; PIDN:
A:Experimental source: strain H37RV
C:Genetics:
A:Gene: RV3512
C:Superfamily: collagen alpha 1(I) chain; fibrillar collagen carboxyl-terminal homology;

Query Match      16.0%; Score 153; DB 2; Length 1079;
Best Local Similarity 43.3%; Pred. No. 0.0003;
Matches 39; Conservative 2; Mismatches 39; Indels 10; Gaps 2;

QY 3 AEGRGTTGGTGADGPGGPIPDGPGNAGGPGGEAGATGGRGPRG-----AGAAASG 55
    |||||
Db 480 AGTGTGAGTGGAGDGGGGGGGGGGAGGGAGGGGGGGGTTGGGNTGGTAGTAGAAG 539
    |||||

QY 56 PGGGAPR---GPHGGAASGLNGCCRCRGARG 82
    |||||
Db 540 NGGAAGKGGAGGGGGTGGGTGGGGGGAGG 569
    |||||

RESULT 5
CGB01S
collagen alpha 1(I) chain - bovine (tentative sequence) (fragments)
C:Species: Bos primigenius taurus (cattle)
C:Date: 24-Apr-1984 #sequence_revision 31-Dec-1993 #text_change 09-Jul-2004
C:Accession: A91193; A91229; A91387; A91211; A91201; A43048; A02853
R:Rauterberg, J.; Timpi, R.; Furthmayr, H.
Eur. J. Biochem. 27, 231-237, 1972
A:Title: Structural characterization of N-terminal antigenic determinants in calf and hu
A:Reference number: A91193; MUID:72255334; PMID:4115172
A:Accession: A91193
A:Molecule type: protein
A:Residues: 1-19 <RAU>
A:Cross-references: UNIPROT:P02453; UNIPARC:UPI0000173B51
A:Experimental source: skin
A:Note: The epsilon carbon of Lys-9, by homology with the rat alpha 1(I) chain, is conver
R:Pietzek, P.P.; Kuehn, K.
Eur. J. Biochem. 52, 77-82, 1975
A:Title: The covalent structure of collagen: amino-acid sequence of the cyanogen-bromide
A:Reference number: A91229; MUID:76022320; PMID:1164916
A:Accession: A91229
A:Molecule type: protein
A:Residues: 20-145 <FIE>
A:Cross-references: UNIPARC:UPI0000173B52
A:Experimental source: skin
A:Note: Lys-103 is hydroxylated and binds glucosylgalactose
R:Pietzek, P.P.; Wendt, P.; Kell, I.; Kuehn, K.
FEBS Lett. 26, 74-76, 1972
A:Title: The covalent structure of collagen: amino acid sequence of alpha1-CB3 from calf
A:Reference number: A91387; MUID:73049499; PMID:4673951
A:Accession: A91387
A:Molecule type: protein
A:Residues: 146-294 <FI2>
A:Cross-references: UNIPARC:UPI0000173B53
A:Experimental source: skin
R:Pietzek, P.P.; Rexrodt, F.W.; Hopfer, K.E.; Kuehn, K.
Eur. J. Biochem. 38, 396-400, 1973
A:Title: The covalent structure of collagen. 2. The amino-acid sequence of alpha1-CB7 fr
A:Reference number: A91211; MUID:74086118; PMID:4359390
A:Accession: A91211
A:Molecule type: protein
A:Residues: 295-562 <FI3>
A:Cross-references: UNIPARC:UPI0000173B54
A:Experimental source: skin
R:Wendt, P.; Mark, K.V.D.; Rexrodt, F.; Kuehn, K.
Eur. J. Biochem. 30, 169-183, 1972
A:Title: The covalent structure of collagen. The amino-acid sequence of the 112 residues
A:Reference number: A91201; MUID:73042276; PMID:4343808
A:Accession: A91201
A:Molecule type: protein
A:Residues: 563-675 <WEN>
A:Cross-references: UNIPARC:UPI0000173B55
```

Nature 310, 337-340, 1984  
A;Title: Human proalpha1(I) collagen gene structure reveals evolutionary conservation of  
A;Reference number: A93335; MUID:84270697; PMID:6462220  
A;Accession: A93335  
A;Molecule type: DNA  
A;Residues: 1-58, 'Q', 60-181 <CHU>  
A;Cross-references: UNIPARC:UPI0000173B3C; EMBL:X00820; NID:G95657; PIDN:CAA25394.1; PID:  
R;Rossaouw, C.M.S.; Vergeer, W.P.; du Plooy, S.J.; Bernard, M.P.; Ramirez, F.; de Wet, W.  
J. Biol. Chem. 262, 15151-15157, 1987  
A;Title: DNA sequences in the first intron of the human pro-alpha 1(I) collagen gene enha  
A;Reference number: I55254; MUID:88033098; PMID:2822714  
A;Accession: I55254  
A;Status: translation not shown; translated from GB/EMBL/DBJ  
A;Molecule type: DNA  
A;Residues: 1-45 <ROS>  
A;Cross-references: UNIPARC:UPI000016A6B2; GB:J02823; NID:G180387; PIDN:AAA51993.1; PID:G  
R;Bornstein, P.; McKay, J.; Morishima, J.K.; Devarayalu, S.; Gellinas, R.E.  
Proc. Natl. Acad. Sci. U.S.A. 84, 8869-8873, 1987  
A;Title: Regulatory elements in the first intron contribute to transcriptional control of  
A;Reference number: A39943; MUID:88097389; PMID:3480516  
A;Accession: A39943  
A;Molecule type: DNA  
A;Residues: 1-34 <BOR>  
A;Cross-references: UNIPARC:UPI000016A6B1; GB:J03559; NID:G180876; PIDN:AAA52052.1; PID:G  
R;Chu, M.L.; de Wet, W.; Bernard, M.; Ramirez, F.  
J. Biol. Chem. 260, 2315-2320, 1985  
A;Title: Fine structural analysis of the human pro-alpha 1 (I) collagen gene. Promoter st  
A;Reference number: I55237; MUID:85130970; PMID:2857713  
A;Accession: I55237  
A;Status: translation not shown; translated from GB/EMBL/DBJ  
A;Molecule type: DNA  
A;Residues: 1-34 <CH2>  
A;Cross-references: UNIPARC:UPI000016A6B1; GB:M10627; NID:G180383; PIDN:AAA51992.1; PID:G  
R;Wirtz, M.K.; Keene, D.R.; Hori, H.; Glanville, R.W.; Steinmann, B.; Rao, V.H.; Holliste  
J. Biol. Chem. 265, 6312-6317, 1990  
A;Title: In vivo and in vitro noncovalent association of excised alpha1(I) amino-terminal  
ome, type VII.  
A;Reference number: A35233; MUID:90202908; PMID:2318855  
A;Accession: A35233  
A;Molecule type: protein  
A;Residues: 33-52 <WIR>  
A;Cross-references: UNIPARC:UPI0000173B3D  
A;Note: This propeptide fragment remained non-covalently bound to a defective, uncleaved  
R;Weil, D.; d'Allesio, M.; Ramirez, F.; de Wet, W.; Coie, W.G.; Chan, D.; Bateman, J.F.  
EMBO J. 8, 1705-1710, 1989  
A;Title: A base substitution in the exon of a collagen gene causes alternative splicing  
A;Reference number: S09400; MUID:89356643; PMID:2767050  
A;Accession: S09400  
A;Molecule type: mRNA  
A;Residues: 156-183 <WEI>  
A;Cross-references: UNIPARC:UPI0000173B3E  
R;Click, E.M.; Bornstein, P.  
Biochemistry 9, 4699-4706, 1970  
A;Title: Isolation and characterization of the cyanogen bromide peptides from the alpha  
A;Reference number: A90567; MUID:71038625; PMID:5529814  
A;Contents: CNBr0-1, CNBr2, CNBr4, CNBr5  
A;Accession: B90567  
A;Molecule type: protein  
A;Residues: 162-198, 'Z', 200-201, 'Z', 203-206, 'Z', 208-209, 'Z', 211-228, 'B', 230, 'BB', 233, 'Z',  
A;Cross-references: UNIPARC:UPI0000173B3F  
A;Experimental source: skin  
A;Note: evidence for 170-allysine  
R;Baetge, B.; Norbom, H.; Diebold  
Eur. J. Biochem. 192, 153-159, 1990  
A;Title: A critical crosslink region in human-bone-derived collagen type I. Specific clea  
A;Reference number: S11372; MUID:90382436; PMID:2169412  
A;Accession: S11372  
A;Molecule type: protein  
A;Residues: 175-187; 274-287, 'P', 289 <BAE>  
A;Cross-references: UNIPARC:UPI0000173B40; UNIPARC:UPI0000173B41  
A;Note: sequence of collagen alpha 1(S)(I) isolated from bone after pepsin digestion  
R;Deak, S.B.; Scholz, P.M.; Amenta, P.S.; Constantinou, C.D.; Levi-Minzi, S.A.; Gonzalez-  
J. Biol. Chem. 266, 21827-21832, 1991

A;Title: The substitution of arginine for glycine 85 of the alpha 1(I) procollagen chain operative melting of intact type I collagen.  
A;Reference number: I55342; MUID:92042092; PMID:1718984  
A;Accession: I55342  
A;Status: translated from GB/EMBL/DBJ  
A;Molecule type: mRNA  
A;Residues: 258-268;1347-1357 <DEA>  
A;Cross-references: UNIPARC:UPI0000173B42; UNIPARC:UPI0000173B43; GB:S67495; NID:g239007  
A;Note: sequences from the 5' and 3' ends only are shown; mutant sequence 263-Arg report R;Morgan, P.H.; Jacobs, H.G.; Segrest, J.P.; Cunningham, L.W.  
J. Biol. Chem. 245, 5042-5048, 1970  
A;Title: Comparative study of glycopeptides derived from selected vertebrate collagens.  
A;Reference number: A92069; MUID:71001508; PMID:4319110  
A;Accession: A92069  
A;Molecule type: protein  
A;Residues: 263-268 <MOR>  
A;Cross-references: UNIPARC:UPI000014DF11  
A;Experimental source: skin  
A;Note: attachment of 2-O-alpha-D-glucosyl-O-beta-D-galactose to 5-hydroxylysine R;Labhard, M.E.; Hollister, D.W.  
Matrix 10, 124-130, 1990  
A;Title: Segmental amplification of the entire helical and telopeptide regions of the collagen type I alpha 1(I) chain  
A;Reference number: S15989; MUID:90326017; PMID:2374517  
A;Accession: S15989  
A;Molecule type: mRNA  
A;Residues: 281-302;402-420;823-843;925-944;1026-1045;1143-1162 <LAB>  
A;Cross-references: UNIPARC:UPI0000173B44; UNIPARC:UPI0000173B45; UNIPARC:UPI0000173B46; R;Wirtz, M.K.; Rao, V.H.; Glanville, R.W.; Labhard, M.E.; Pretorius, P.J.; de Vries, W.N.  
Connect. Tissue Res. 29, 1-11, 1993  
A;Title: A cysteine for glycine substitution at position 175 in an alpha 1 (I) chain of collagen type I  
A;Reference number: I52905; MUID:93339042; PMID:8339541  
A;Accession: I52905  
A;Status: translated from GB/EMBL/DBJ  
A;Molecule type: mRNA  
A;Residues: 342-352; 'C', 354-359 <WID>  
A;Cross-references: UNIPARC:UPI000070EB6; GB:S64717; NID:g408195; PIDN:AAB27677.1; PID: A;Note: mutant sequence from patient with osteogenesis imperfecta R;Bernard, M.P.; Chu, M.L.; Myers, J.C.; Ramirez, F.; Eikenberry, E.F.; Prockop, D.J.  
Biochemistry 22, 5213-5223, 1983  
A;Title: Nucleotide sequences of complementary deoxyribonucleic acids for the proalpha 1(I) chain of collagen type I  
A;Reference number: A90476; MUID:84080385; PMID:6689127  
A;Accession: A90476  
A;Molecule type: mRNA  
A;Residues: 425-1250, 'S', 1330-1390, 'X', 1392-1464 <BER>  
A;Cross-references: UNIPARC:UPI0000173B4A; GB:K01228; NID:g180391; PIDN:AAA51995.1; PID: A;Note: sequence partially completed for missing nucleotides by A29439 R;Chu, M.L.; Gargiulo, V.; Williams, C.J.; Ramirez, F.  
J. Biol. Chem. 260, 691-694, 1985  
A;Title: Multixon deletion in an osteogenesis imperfecta variant with increased type I collagen  
A;Reference number: A22161; MUID:85104934; PMID:2981843  
A;Accession: A22161  
A;Molecule type: DNA  
A;Residues: 472-594, 'R', 596-607 <CH3>  
A;Cross-references: UNIPARC:UPI000011F786; GB:K03178; GB:K03179; NID:g179612; NID:g179613  
A;Note: the authors translated the codon CGT for residue 595 as Pro R;Wallis, G.A.; Starman, B.J.; Zinn, A.B.; Byers, P.H.  
Am. J. Hum. Genet. 46, 1034-1040, 1990  
A;Title: Variable expression of osteogenesis imperfecta in a nuclear family is explained by a mutation in the alpha 1(I) chain of collagen type I  
A;Reference number: A35336; MUID:90252792; PMID:2339700  
A;Accession: A35336  
A;Molecule type: mRNA  
A;Residues: 710-720, 'E', 722-737, 'E', 739-745 <WAL>  
A;Cross-references: UNIPARC:UPI0000173B4B  
A;Note: the authors translated the codons CAG for 721 and CGT for 738 as Glu R;Forlino, A.; Zolezzi, F.; Valli, M.; Pignatti, P.F.; Cetta, G.; Brunelli, P.C.; Mottes  
Hum. Mol. Genet. 3, 2201-2206, 1994  
A;Title: Severe (type III) osteogenesis imperfecta due to glycine substitutions in the alpha 1(I) chain of collagen type I  
A;Reference number: I54365; MUID:95187161; PMID:7881420  
A;Accession: I54365  
A;Status: translated from GB/EMBL/DBJ  
A;Molecule type: DNA  
A;Residues: 746-766, 'S', 768-781 <FOR>

A;Cross-references: UNIPARC:UPI000016AGFA; GB:L47667; NID:g1009093; PIDN:AAB59576.1; PID: R;Chessler, S.D.; Wallis, G.A.; Byers, P.H.  
J. Biol. Chem. 268, 18218-18225, 1993  
A;Title: Mutations in the carboxyl-terminal propeptide of the pro alpha 1(I) chain of type I collagen  
A;Reference number: A47426; MUID:93352846; PMID:8349697  
A;Accession: A47426  
A;Molecule type: mRNA  
A;Residues: 1179-1276, 'H', 1278-1336, 1339-1387, 'R', 1389-1464 <CHE>  
A;Cross-references: UNIPARC:UPI0000073A2A; GB:S64596; NID:g407589; PIDN:AAB27856.1; PID: A;Note: sequence extracted from NCBI backbone (NCBIN:136444, NCBIP:136445)  
A;Note: does not represent an experimentally determined sequence but three different mutations  
A;Accession: B47426  
A;Molecule type: mRNA  
A;Residues: 1179-1464 <CH4>  
A;Cross-references: UNIPARC:UPI0000173B4C  
A;Experimental source: normal dermal fibroblast culture  
A;Accession: C47426  
A;Molecule type: mRNA  
A;Residues: 1179-1276, 'H', 1278-1464 <CH5>  
A;Cross-references: UNIPARC:UPI0000173B4D  
A;Experimental source: fetal cell 86-237  
A;Accession: D47426  
A;Molecule type: mRNA  
A;Residues: 1179-1336, 1339-1464 <CH6>  
A;Cross-references: UNIPARC:UPI0000173B4E  
A;Experimental source: fetal cell 86-146  
A;Accession: E47426  
A;Molecule type: mRNA  
A;Residues: 1179-1387, 'R', 1389-1464 <CH7>  
A;Cross-references: UNIPARC:UPI0000173B4F  
A;Experimental source: fetal cell 88-251  
R;Cohn, D.H.; Apone, S.; Byre, D.R.; Starman, B.J.; Andreassen, P.; Charbonneau, H.; Nict  
J. Biol. Chem. 263, 14605-14607, 1988  
Query Match 15.8%; Score 151.5; DB 1; Length 1464;  
Best Local Similarity 37.2%; Pred. No. 0.00051;  
Matches 48; Conservative 42; Mismatches 42; Indels 35; Gaps 8;  
QY 3 ASGR-GTGGSTGDA--DGGPGGPIPDGPGNAGGPGGAGATGGRGPRGA---GAARASG 55  
Db 825 AKGPGDAGKADGADGPPGPPAGPGPIGNVGAQKAGSAGSGPPGATGPGGAAGRVG 884  
QY 56 P---GGGAPRPHGGAASGLNGCCRCGARGPESRLLEFLYPMPFATPMEALARRSLAQ 111  
Db 885 PPGPSGNAGPPGPPGPPAG-----KEGKGPRGE-----TGPAGRPGEV----- 922  
QY 112 DAPPLPVP 120  
Db 923 -GPPGP-PG 929  
RESULT 7  
D70807  
hypothetical glycine-rich protein RV3514 - Mycobacterium tuberculosis (strain H37RV)  
C;Species: Mycobacterium tuberculosis  
C;Date: 17-Jul-1998 #sequence\_revision 17-Jul-1998 #text\_change 20-Jun-2000  
C;Accession: D70807  
R;Cole, S.T.; Brosch, R.; Parkhill, J.; Garnier, T.; Churcher, C.; Harris, D.; Gordon, S.; Connor, R.; Davies, R.; Devlin, K.; Feldwell, T.; Gentles, S.; Hamlin, N.; Holroyd, S.; Rajandream, M.A.; Rogers, J.; Rutter, S.; Seeger, K.; Skelton, S.; Squares, S.  
Nature 393, 537-544, 1998  
A;Authors: Sgares, R.; Sulston, J.E.; Taylor, K.; Whitehead, S.; Barrell, B.G.  
A;Title: Deciphering the biology of Mycobacterium tuberculosis from the complete genome  
A;Reference number: A70500; MUID:98295987; PMID:9634230  
A;Accession: D70807  
A;Status: preliminary; nucleic acid sequence not shown; translation not shown  
A;Molecule type: DNA  
A;Residues: 1-1489 <COL>  
A;Cross-references: UNIPARC:UPI00000D3ABE; GB:AL022022; GB:AL123456; NID:g3261554; PIDN: A;Experimental source: strain H37RV  
C;Genetics:  
A;Gene: RV3514  
C;Superfamily: collagen alpha 1(IV) chain

Query Match      15.7%; Score 151; DB 2; Length 1489;  
Best Local Similarity 41.1%; Pred. No. 0.00057;  
Matches 37; Conservative 2; Mismatches 31; Indels 20; Gaps 3;

OY    5 GRGTGSGTGDADGGGP-----CIPDPPGNAGPGGEA-----GATGGRGPRCA 48  
         | : |||||  
DB    140 GAGAACAPQCGPFGGAGLWNGNPGGAGSGGGTGGAGGAGGWLFVGVGAGGVGGAGG 199

OY    49 GAARASGPG----GGAPRPHGGAASGLNG 74  
         | : |||||  
DB    200 GTGGAGGPGGLIWGGGAGGVGGAGGGTG 229

RESULT 8

F70806  
Hypochemical glycine-rich protein Rv3508 - Mycobacterium tuberculosis (strain H37RV)  
C.Species: Mycobacterium tuberculosis  
C.Date: 17-Jul-1998 #sequence\_revision 17-Jul-1998 #text\_change 09-Jul-2004  
C.Accession: F70806  
R.Cole, S.T.; Brosch, R.; Parkhill, J.; Garnier, T.; Churcher, C.; Harris, D.; Gordon, S.; Connor, R.; Davies, R.; Devlin, K.; Fellwell, T.; Gentles, S.; Hamlin, N.; Holroyd, S.; Rajandream, M.A.; Rogers, J.; Rutter, S.; Seeger, K.; Skelton, S.; Squares, S.  
Nature 393, 537-544, 1998  
A.Authors: Sqares, R.; Sulston, J.E.; Taylor, K.; Whitehead, S.; Barrell, B.G.  
A>Title: Deciphering the biology of Mycobacterium tuberculosis from the complete genome  
A.Reference number: A70500; UID:98295987; PMID:9634230  
A.Accession: F70806  
A.Status: preliminary; nucleic acid sequence not shown; translation not shown  
A.Molecule type: DNA  
A.Residues: 1-1901 <COL>  
A.Cross-references: UNIPROT:O53553; UNIPARC:UI000013C2A6; GB:AL022022; GB:AL123456; MID  
A.Experimental source: strain H37rv  
C.Genetics:  
A.Gene: Rv3508  
C.Superfamily: collagen alpha 1(IV) chain

```

Query Match      15.7%; Score 151; DB 2; Length 1901;
Best Local Similarity 41.1%; Pred. No. 0.00071;
Matches 37; Conservative 2; Mismatches 31; Indels 20; Gaps 3;

QY  5  GRGTGSGTGADCGPGGP-----GIPDGPNGNAGGPGEA-----GATGGRGPRGA 48
DB  140 GAGAAGAPGQAGCGPGGAGLWNGNGPGGAGSGGGTGGAGGAGGWLFGVGGAGGVGGAGG 199

QY  49  GAARASGPG-----GGAPRGPHPHGGAAASGLNG 74
DB  200 GTGGAGCGPGLIWWGGGAGGVGGAGCGGTGG 229

RESULT 9
T32812
hypothetical protein H17B01.2 - Caenorhabditis elegans
C:Species: Caenorhabditis elegans
C:Date: 29-Oct-1999 #sequence_revision 29-Oct-1999 #text_change 09-Jul-2004
C:Accession: T32812
R:Gattung, S.; Maggi, L.
submitted to the EMBL Data Library, December 1997
A:Description: The sequence of C. elegans cosmid H17B01.
A:Reference number: Z21227
A:Accession: T32812
A:Status: preliminary; translated from GB/EMBL/DBJ
A:Molecule type: DNA
A:Residues: 1-530 <GAT>
A:Cross-references: UNIPROT:O61209; UNIPARC:UPI00000792BA; EMBL:AF040646; PIDN:AAB94986.
A:Experimental source: strain Bristol N2; clone H17B01
C:Genetics:
A:Gene: CESP:H17B01.2
A:Map position: 2
A:Introns: 42/3; 58/1; 173/3; 268/2; 308/2; 340/1; 364/2; 387/3

Query Match      15.7%; Score 150.5; DB 2; Length 530;
Best Local Similarity 32.3%; Pred. No. 0.00024;

```

Matches	51; Conservative	9; Mismatches	65; Indels	33; Gaps	8;	
Qy	5	GRGTGGSTGDADGGGPGIPDGGNAGGPGGAGATGGRG-PRGAGAAASGPGGGAPRG	63			
Db	96	GLGTPGGPGNPGGGLGTPGPG-GTTPGGPNPGGLGGPGNPGGPGTGGTGGTGGQPHG	154			
Qy	64	PHGGAASGL---NG-----CCRC-----GARGPESRLLEFYLAHPF---ATPMEAELAR	106			
Db	155	TAGGTYYVIIVQNGQRVIVNCATRLCCVGRNITYSVVTDSPVPPVPLVVTGTP-----	207			
Qy	107	RLSAQDAPPLPV-----PGVLLKEFTVSGNILLTIRLT	138			
Db	208	-SIVTVTPSPSVTPGGRPGSTTPPATFPFGSVSTVRYT	244			
RESULT	10					
H70589						
		hypothetical glycine-rich protein Rv2853 - Mycobacterium tuberculosis (strain H37RV)				
C:Species:		Mycobacterium tuberculosis				
C:Date:		17-Jul-1998 #sequence_revision 17-Jul-1998 #text_change 31-Mar-2003				
C:Accession:		H70589				
		R:Cole, S.T.; Brosch, R.; Parkhill, J.; Garnier, T.; Churcher, C.; Harris, D.; Gordon, S.; Connor, R.; Davies, R.; Devlin, K.; Felwell, T.; Gentles, S.; Hamlin, N.; Holroyd, S.; Nardaram, M.A.; Rogers, J.; Rutter, S.; Seeger, K.; Skelton, S.; Squares, S.; Nature 393, 537-544, 1998				
A:Authors:		Squires, R.; Sulston, J.E.; Taylor, K.; Whitehead, S.; Barrall, B.G.				
A:Title:		Deciphering the biology of Mycobacterium tuberculosis from the complete genome				
A:Reference number:		A70500; MUID:98295987; PMID:9634230				
A:Accession:		H70589				
A:Status:		preliminary; nucleic acid sequence not shown; translation not shown				
A:Molecule type:		DNA				
A:Residues:		1-615 <COL>				
A:Cross-references:		UNIPARC:UPI00000D3B10; GB:Z95207; GB:AL123456; NID:G3261745; PIDN:CAI				
A:Experimental source:		strain H37RV				
C:Genetics:						
A:Gene:		Rv2853				
C:Superfamily:		uncharacterized glycine-rich protein, PE motif containing				
Query Match		15.6%; Score 149.5; DB 2; Length 615;				
Best Local Similarity		49.3%; Pred. No. 0.00033;				
Matches		35; Conservative	5; Mismatches	26; Indels	5; Gaps	3;
Qy	7	GTGGSTGDADGGGPGIPDGGNAGG--PGEAGATGGRGPRGA-GAARASGPGGGAPRG	63			
Db	532	GTGGNGDGGAGAPGL--GGAGNGGWLIIQSGSTGGGAGGAGGAGGSGGAG	589			
Qy	64	PHGGAASGLNG	74			
Db	590	GAGDTTSGNG	600			

RESULT 11  
T32734  
myosin-IA - Acanthamoeba castellanii  
C:Species: Acanthamoeba castellanii  
C:Date: 29-Oct-1999 #sequence\_revision 29-Oct-1999 #text\_change 09-Jul-2004  
C:Accession: T32734  
R:Lee, W.L.; Ostap, E.M.; Zot, H.G.; Pollard, T.D.  
submitted to the EMBL Data Library, August 1998  
A:Description: Hydrodynamic and ligand binding properties of Acanthamoeba myosin-IA GPA/1  
A:Reference number: Z21216  
A:Accession: T32734  
A:Status: preliminary; translated from GB/EMBL/DBJ  
A:Molecule type: DNA  
A:Residues: 1-1215 <LEE>  
A:Cross-references: UNIPROT:O77202; UNIPARC:UPI000007E5D0; EMBL:AF085185; NID:g3599477; I  
A:Experimental source: strain Neff  
C:Genetics:  
A:Gene: M1A  
A:Introns: 1/3; 41/3; 72/2; 103/2; 162/3; 184/1; 217/1; 296/1; 340/3; 390/3; 447/3; 500/3  
C:Superfamily: protozoan myosin heavy chain 1B; myosin motor domain homology; SH3 homolog  
F:164-674/Domain: myosin motor domain homology <MNO>



Matches	54;	Conservative	5;	Mismatches	50;	Indels	75;	Gaps	7;
Qy	6	RGTGGSTGDADGPGPGIP	-----	DPGGNAGGPGGAGATGGRGPRG	-	47			
Db	871	KDRGETGPAGPGAPGAPGAPGPGVGPAGKNGDRGETGPAGPGAGPGAGPGPGP	930						
Qy	48	-----	AGAAASGPGGAPRPHGGAAS	-	70				
Db	931	RGDKETGEQDRGMKGHRGFSGLQGPFGPPGAPGEOGPGSAGSPAG	-	PRGPPGSAGAA	988				
Qy	71	---GLNCC-CRCGARGPESRLLEFYLLAMPFATPMEAEIARRSLAQDAPPLPVGVLKEF	126						
Db	989	GKDLNGLPGPIGPPGPRGTGEVGVGPPGP	-----	GPPGP-EGPPSGGF	1034				
Qy	127	TVSG	130						
Db	1035	DLDG	1038						

Search completed: February 4, 2006, 12:18:10  
Job time : 42 secs

**This Page Blank (uspto)**

GenCore version 5.1.7  
Copyright (c) 1993 - 2006 Bioceleration Ltd.

OM protein - protein search, using sw model

Run on: February 4, 2006, 12:10:02 ; Search time 248 Seconds  
(without alignments)  
512.077 Million cell updates/sec

Title: US-09-529-206E-4

Perfect score: 959

Sequence: 1 MQAEGRGTCGTGADPGG.....WITQCFLPVFLAQPSPGQR 180

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 2166443 seqs, 705528306 residues

Total number of hits satisfying chosen parameters: 2166443

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

UniProt\_05.80.\*

1: uniprot\_sprot.\*

2: uniprot\_trembl.\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

#### SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	959	100.0	180	1 CTG1B_HUMAN	P78358 homo sapien
2	959	100.0	180	2 Q7LBY4_HUMAN	Q7LBY4 homo sapien
3	595	62.0	210	1 CTAG2_HUMAN	O75638 homo sapien
4	581	60.6	142	2 Q9NY13_HUMAN	Q9NY13 homo sapien
5	222	23.1	143	1 ITBA2_HUMAN	Q14657 homo sapien
6	222	23.1	143	2 Q5HY39_HUMAN	Q5HY39 homo sapien
7	174.5	18.2	627	1 HYDL_STRCO	P42534 streptomyce
8	172.5	18.0	563	2 Q6J6N0_9ARAC	O6J6N0 araneus ven
9	169	17.6	617	2 Q5B4Q0_EMENI	O5B4Q0 aspergillus
10	168.5	17.6	1002	2 Q9BIU8_ARGTR	Q9BIU8 argiope tri
11	167.5	17.5	922	2 Q44367_MYTGD	Q44367 mytilus edu
12	165.5	17.3	922	2 Q8MW53_MYTGA	Q8MW53 mytilus gal
13	162	16.9	420	2 Q6RF45_ARATH	O6RF45 arabidopsis
14	162	16.9	651	2 Q9BIU9_ARGTR	Q9BIU9 argiope tri
15	161	16.8	524	2 Q8I7U1_9ARAC	O8I7U1 araneus ven
16	160.5	16.7	813	2 Q636W4_BACCZ	O636W4 bacillus ce
17	160.5	16.7	1168	2 Q6HF99_BACHK	Q6HF99 bacillus th
18	160.5	16.7	1408	2 Q7U022_MYCBL	Q7U022 mycobacteri
19	160.5	16.7	2249	2 Q9NHW4_NEPCL	Q9NHW4 nephila cla
20	160	16.7	604	2 Q9L252_STRCO	Q9L252 streptomyce
21	160	16.7	871	2 Q4358_NEPCL	O4358 nephila cla
22	159.5	16.6	1884	2 Q9NHW2_9ARAC	Q9NHW2 nephila ina
23	159	16.6	904	2 Q76271_MYTGD	Q76271 mytilus edu
24	158.5	16.5	678	2 Q93486_ONCMY	Q93486 oncorhynch
25	158.5	16.5	1458	2 Q910B9_ONCMY	Q910B9 oncorhynch
26	157.5	16.4	597	2 Q58MV1_9CAUD	O58MV1 cyanophage
27	157	16.4	410	2 Q16988_ARADI	Q16988 araneus dia
28	157	16.4	626	2 Q9NHW1_9ARAC	Q9NHW1 nephila ina
29	156.5	16.3	1046	1 IF2_STRAW	Q82K53 streptomyce
30	155.5	16.2	145	2 Q61351_DROME	O61351 drosophila
31	155.5	16.2	641	1 EBN1_EBV	P03211 epstein-bar

32	155.5	16.2	641	2 Q777E1_9GAMA	Q777E1 human herpe
33	155	16.2	894	2 Q8MW54_MYTGA	Q8MW54 mytilus gal
34	155	16.2	902	2 O16161_MYTGD	O16161 mytilus edu
35	154.5	16.1	478	2 Q8LWX2_BACAN	Q8LWX2 bacillus an
36	154.5	16.1	481	2 Q6HV21_BACAN	Q6HV21 bacillus an
37	154.5	16.1	905	2 Q8MW55_MYTGA	Q8MW55 mytilus gal
38	154.5	16.1	1953	2 Q9BIT7_9ARAC	Q9BIT7 mycobacteri
39	153	16.0	1079	2 Q6MW7 MYCTU	Q6MW7 mycobacteri
40	153	16.0	1449	2 Q6NZ15_BRARE	Q6NZ15 brachydanio
41	153	16.0	1449	2 Q6PE19_BRARE	Q6PE19 brachydanio
42	152.5	15.9	734	2 Q6IVJ4_CTOIN	Q6IVJ4 ciona intes
43	152.5	15.9	779	1 COIAl_BOVIN	P02453 bos taurus
44	152.5	15.9	1938	2 Q7TWC0_MYCBO	Q7TWC0 mycobacteri
45	152	15.8	219	2 Q9D9S4_MOUSE	Q9D9S4 mus musculu

#### ALIGNMENTS

RESULT 1				
ID	CTG1B_HUMAN	STANDARD;	PRT;	180 AA.
AC	P78358;			
DT	15-DEC-1998 (Rel. 37, Created)			
DT	15-DEC-1998 (Rel. 37, Last sequence update)			
DT	10-MAY-2005 (Rel. 47, Last annotation update)			
DE	Cancer/testis antigen 1B (Autoimmunogenic cancer/testis antigen NY-ESO-1).			
GN	Name=CTAG1B; Synonym=CTAG, CTAG1;			
OS	Homo sapiens (Human).			
OC	Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;			
OC	Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini; Hominidae;			
OC	Homo.			
OX	NCBI_TaxID=9606;			
RN	[1]			
RP	NUCLEOTIDE SEQUENCE.			
RX	MEDLINE=97203161; PubMed=9050879; DOI=10.1073/pnas.94.5.1914;			
RA	Chen Y.-T., Scanlan M.J., Sahin U., Tuerki O., Gure A.O., Tsang S., Williamson B., Stockert E., Pfreundschuh M., Old L.J.;			
RT	"A testicular antigen aberrantly expressed in human cancers detected by autologous antibody screening."			
RL	Proc. Natl. Acad. Sci. U.S.A. 94:1914-1918(1997).			
RN	[2]			
RP	NUCLEOTIDE SEQUENCE.			
RC	TISSUE=Melanoma;			
RX	MEDLINE=98289662; PubMed=9626360;			
RA	Lethe B., Lucas S., Michaux L., de Smet C., Godelaine D., Serrano A., de Plaen E., Boon T.;			
RT	"LAGE-1, a new gene with tumor specificity."			
RL	Int. J. Cancer 76:903-908(1998).			
RN	[3]			
RP	NUCLEOTIDE SEQUENCE.			
RX	MEDLINE=98430882; PubMed=9759882;			
RA	Wang R.-F., Johnston S.L., Zeng G., Topalian S.L., Schwartzentruber D.J., Rosenberg S.A.;			
RT	"A breast and melanoma-shared tumor antigen: T cell responses to antigenic peptides translated from different open reading frames."			
RL	J. Immunol. 161:3596-3606(1998).			
CC	-!- TISSUE SPECIFICITY: Expressed in testis and ovary and in a wide variety of cancers. Detected in uterine myometrium.			
CC	-!- SIMILARITY: Belongs to the CTAG family.			
CC	This Swiss-Prot entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL outstation - the European Bioinformatics Institute. There are no restrictions on its use as long as its content is in no way modified and this statement is not removed.			
DR	EMBL; U87459; AAB49693.1; -; mRNA.			
DR	EMBL; AJ003149; CA05908.1; -; mRNA.			
DR	EMBL; AF038567; AAD05202.1; -; mRNA.			
DR	FGNC; HGNC:2491; CTAG1B.			
DR	MIM; 300156; -.			

```
KW Antigen; Transmembrane.
FT TRANSMEM 156 172 Potential.
FT COMPIAS 5 82 Gly-rich.
SC SEQUENCE 180 AA; 17992 MW; 5122C5C2C8BE1569 CRC64;

Query Match
Best Local Similarity 100.0%; Score 959; DB 1; Length 180;
Matches 180; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 MQAEGRTGSGTGDADPGGPGIPDPGPGNAGGPGAGATGGRPGAGAAASGPGGA 60
DB 1 MQAEGRTGSGTGDADPGGPGIPDPGPGNAGGPGAGATGGRPGAGAAASGPGGA 60
QY 61 PRGPHGGAASGLNGCCRCGARGPESRLLEFYLAAMPATPMEAEARRSLAQDAPLPVPG 120
DB 61 PRGPHGGAASGLNGCCRCGARGPESRLLEFYLAAMPATPMEAEARRSLAQDAPLPVPG 120
QY 121 VLLKEFTVSGNLTIRLTAADHRQLQLSISSCLQLSLMLWITQCFLPVFLAQPSPGQRR 180
DB 121 VLLKEFTVSGNLTIRLTAADHRQLQLSISSCLQLSLMLWITQCFLPVFLAQPSPGQRR 180

RESULT 2
ID Q7LBY4 HUMAN PRELIMINARY; PRT; 180 AA.
AC Q7LBY4;
DT 10-MAY-2005 (T-EMBLrel. 30, Created)
DI 10-MAY-2005 (T-EMBLrel. 30, Last sequence update)
DE Cancer/testis antigen 1-B (NY-ESO-1/LAGE-2a protein) (Cancer/testis
DE antigen 1-A).
GN Names=CTAG1-B; Synonyms=CTAG1-A, LAGE-2;
OS Homo sapiens (Human)
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini; Hominidae;
OC Homo.
OX NCBI_TaxID=9606;
RN [1]
RP NUCLEOTIDE SEQUENCE.
RX MEDLINE=21303288;
RA Galgoczy P., Rosenthal A., Platzer M.;
RT "Human-mouse comparative sequence analysis of the NPMO gene reveals an
RT alternative promoter within the neighboring G6PD gene.";
RL Gene 271:93-98 (2001).
RN [2]
RP NUCLEOTIDE SEQUENCE.
EX MEDLINE=21566154; PubMed=11709543; DOI=10.1093/hmg/10.22.2557;
FA Aradhya S., Bardaro T., Galgoczy P., Yamagata T., Eposito T.,
RA Patlan H., Ciccodicola A., Kenwick S., Platzer M., D'Urso M.,
RA Nelson D.L.;
PT "Multiple pathogenic and benign genomic rearrangements occur at a 35
PT kb duplication involving the NEMO and LAGE2 genes.";
RL Hum. Mol. Genet. 10:2557-2567 (2001).
RN [3]
RP NUCLEOTIDE SEQUENCE.
RA Platzer M.;
RP Submitted (MAY-2001) to the EMBL/GenBank/DBJ databases.
RN [4]
RP NUCLEOTIDE SEQUENCE.
RA Platzer M.;
RP Submitted (JUL-2001) to the EMBL/GenBank/DBJ databases.
RN [5]
RP NUCLEOTIDE SEQUENCE.
RA Galgoczy P., Platzer M.;
RP Submitted (MAR-2003) to the EMBL/GenBank/DBJ databases.
RN [6]
RP NUCLEOTIDE SEQUENCE.
EX MEDLINE=99454989; PubMed=10523621;
RA De Smet C., Lurquin C., Lethe B., Martelange V., Boon T.;
RT "DNA methylation is the primary silencing mechanism for a set of germ
RT line- and tumor-specific genes with a CpG-rich promoter.";
RL Mol. Cell. Biol. 19:7327-7335 (1999).
RN [7]
RP NUCLEOTIDE SEQUENCE.
RA Lethe B.G.;
RL Submitted (MAR-2000) to the EMBL/GenBank/DBJ databases.
DR EMBL; AF277315; AAL27014.1; -; Genomic_DNA.
DR EMBL; AJ275977; CAB76943.1; -; Genomic_DNA.
DR EMBL; AF277315; AAL27013.1; -; Genomic_DNA.
SQ SEQUENCE 180 AA; 17992 MW; 5122C5C2C8BE1569 CRC64;

Query Match
Best Local Similarity 100.0%; Score 959; DB 2; Length 180;
Matches 180; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 MQAEGRTGSGTGDADPGGPGIPDPGPGNAGGPGAGATGGRPGAGAAASGPGGA 60
DB 1 MQAEGRTGSGTGDADPGGPGIPDPGPGNAGGPGAGATGGRPGAGAAASGPGGA 60
QY 61 PRGPHGGAASGLNGCCRCGARGPESRLLEFYLAAMPATPMEAEARRSLAQDAPLPVPG 120
DB 61 PRGPHGGAASGLNGCCRCGARGPESRLLEFYLAAMPATPMEAEARRSLAQDAPLPVPG 120
QY 121 VLLKEFTVSGNLTIRLTAADHRQLQLSISSCLQLSLMLWITQCFLPVFLAQPSPGQRR 180
DB 121 VLLKEFTVSGNLTIRLTAADHRQLQLSISSCLQLSLMLWITQCFLPVFLAQPSPGQRR 180

RESULT 3
ID CTAG2 HUMAN STANDARD; PRT; 210 AA.
AC Q75638; O75637; Q9UB00; Q9UJ89; Q9Y479;
DT 30-MAY-2000 (Rel. 39, Created)
DT 30-MAY-2000 (Rel. 39, Last sequence update)
DT 30-MAY-2000 (Rel. 47, Last annotation update)
DE Cancer/testis antigen 2 (LAGE-1 protein) (ESO-2 protein).
GN Name=CTAG2; Synonyms=ESO2, LAGE1;
OS Homo sapiens (Human)
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini; Hominidae;
OC Homo.
OX NCBI_TaxID=9606;
RN [1]
RP NUCLEOTIDE SEQUENCE (ISOFORMS LAGE-1A AND LAGE-1B).
RC TISSUE=Melanoma;
RX MEDLINE=98289662; PubMed=9626360;
RA Lethe B., Lucas S., Michaux L., de Smet C., Godelaine D., Serrano A.,
RA de Plaen E., Boon T.;
RT "LAGE-1, a new gene with tumor specificity.";
RL Int. J. Cancer 76:903-908 (1998).
RN [2]
RP NUCLEOTIDE SEQUENCE (ISOFORMS LAGE-1A AND LAGE-1B), AND VARIANTS
RC GLN-6; GLN-89 AND ARG-138.
RX TISSUE=Melanoma;
RX MEDLINE=99325550; PubMed=10399963;
RA Aarnoudse C.A., Van den Doel P.B., Heemskerk B., Schrier P.I.;
RT "Interleukin-2-induced, melanoma-specific T cells recognize CAMEL, an
RT unexpected translation product of LAGE-1.";
RL Int. J. Cancer 82:442-448 (1999).
RN [3]
RP NUCLEOTIDE SEQUENCE [LARGE SCALE MRNA] (ISOFORM LAGE-1B), AND VARIANT
RC ARG-138.
RX TISSUE=Placenta;
RX MEDLINE=22388257; PubMed=12477932; DOI=10.1073/pnas.242603899;
RA Strausberg R.L., Feingold E.A., Grouse L.H., Derge J.G.,
RA Klausner R.D., Collins F.S., Wagner L., Shenmen C.M., Schuler G.D.,
RA Altschul S.P., Zeeberg B., Buetow K.H., Schaefer C.F., Bhat N.K.,
RA Hopkins R.F., Jordan H., Moore H., Moore A., Rubin G.M., Hong F.,
RA Diachenko L., Marusina K., Farmer A.A., Tashiro H., Scheetz T.E.,
RA Brownstein M.J., Usdin T.B., Toshiyuki S., Carninci P., Prange C.,
RA Stapleton M., Soares M.B., Bonaldo M.F., Casavant T.L., Abramson R.D.,
RA Raha S.S., Loquellano N.A., Peters G.J., Abramson R.D., Mullany S.J.,
RA Bosak S.A., McEwan P.J., McKernan K.J., Malek J.A., Gunaratne P.H.,
RA Richards S., Worley K.C., Hale S., Garcia A.M., Gay L.J., Hulyk S.W.,
RA Villalón D.K., Muzny D.M., Sodergren E.J., Lu X., Gibbs R.A.,
RA Fahey J., Helton E., Kettman M., Madan A., Rodriguez S., Sanchez A.,
```

RA Whiting M., Madan A., Young A.C., Shevchenko Y., Bouffard G.G.,  
RA Blakesley R.W., Touchman J.W., Green E.D., Dickson M.C.,  
RA Rodriguez A.C., Grimwood J., Schmutz J., Myers R.M.,  
RA Butterfield Y.S.N., Krzywinski M.I., Skalska U., Smalilus D.E.,  
RA Schnerch A., Schein J.E., Jones S.J.M., Marra M.A.;  
RT "Generation and initial analysis of more than 15,000 full-length human  
RT and mouse cDNA sequences.";  
RL Proc. Natl. Acad. Sci. U.S.A. 99:16899-16903(2002).  
CC -!- ALTERNATIVE PRODUCTS:  
CC Event=Alternative splicing; Named isoforms=2;  
CC Name=LAGE-1B; Synonyms=LAGE-1L;  
CC IsoId=O75638-1; Sequence=Displayed;  
CC Name=LAGE-1A; Synonyms=LAGE-1S;  
CC IsoId=O75638-2; Sequence=VSP\_004301;  
CC -!- TISSUE SPECIFICITY: Testis and very low level in placenta and in  
CC some uterus samples. Observed in 25-50% of tumor samples of  
CC melanomas, non-small-cell lung carcinomas, bladder, prostate and  
CC head and neck cancers.  
CC -!- DOMAIN: A transmembrane domain is present in isoform LAGE-1A.  
CC -!- SIMILARITY: Belongs to the CTAG family.  
CC -----  
CC This Swiss-Prot entry is copyright. It is produced through a collaboration  
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
CC the European Bioinformatics Institute. There are no restrictions on its  
CC use as long as its content is in no way modified and this statement is not  
CC removed.  
CC -----  
DR EMBL; AJ223093; CAA11117.1; -; Genomic DNA.  
DR EMBL; AJ223093; CAA11116.1; -; Genomic DNA.  
DR EMBL; AJ223040; CAA11043.1; -; mRNA.  
DR EMBL; AJ223041; CAA11044.1; -; mRNA.  
DR EMBL; AJ012834; CAA10194.1; -; mRNA.  
DR EMBL; AJ012835; CAA10196.1; -; mRNA.  
DR EMBL; BC002833; AAH02833.1; -; mRNA.  
DR Ensembl; ENSG00000126890; Homo sapiens.  
DR HGNC; HGNC:2492; CTAG2.  
DR H-InvDB; HIX0017163; -.  
DR MIM; 300396; -.  
KW Alternative splicing; Antigen; Polymorphism; Transmembrane.  
FT COMPIAS 5 79 Gly-rich.  
FT COMPIAS 183 188 Poly-Pro.  
FT VARSPPLIC 135 210 MSVDDQDRGAGRMVVGWGLSGASPEGQKARDLRTPKHV  
FT SEQRPTGPPPPGPGGQDGCGRVAFNVMSFAPHI -> IR  
FT LTAADHRQLQSLSSCLQLSLIMWITQCLPLVFLAQPDSG  
FT ORR (in isoform LAGE-1A).  
FT /FTId=VSP\_004301.  
FT R -> Q.  
FT VARIANT 6 6 /FTId=VAR\_007855.  
FT E -> Q.  
FT VARIANT 89 89 /FTId=VAR\_007856.  
FT W -> R.  
FT VARIANT 138 138 /FTId=VAR\_007857.  
SQ SEQUENCE 210 AA; 21120 MW; 8BE0EE00AE55E8BE CRC64;  
  
Query Match 62.0%; Score 595; DB 1; Length 210;  
Best Local Similarity 82.0%; Pred. No. 3.6e-35;  
Matches 109; Conservative 11; Mismatches 13; Indels 0; Gaps 0;  
  
QY 1 MQAEGRTGSGTGDADPGPGPTPDGPGNAGGPGAGATGGRPGAGAGARASGPGGA 60  
DB 1 MQAEGRTGSGTGDADPGPGPTPDGPGNAGGPGAGATGGRPGAGAGARASGPGGA 60  
  
QY 61 PRPGHGAASGLNGCCRCGARGPESRLLEFLYLPAMPATPMEALARRSLAQAADPLPVPG 120  
DB 61 PRPGHGAASGLNGCCRCGARGPESRLLEFLYLPAMPATPMEALARRSLAQAADPLPVPG 120  
  
QY 121 VLKFTVSGNIL 133  
DB 121 AVLKFTVSGNLL 133  
  
RESULT 4  
Q9NY13\_HUMAN

ID Q9NY13\_HUMAN PRELIMINARY; PRT; 142 AA.  
AC Q9NY13;  
DT 01-OCT-2000 (TrEMBLrel. 15, Created)  
DT 01-OCT-2000 (TrEMBLrel. 15, Last sequence update)  
DT 01-OCT-2002 (TrEMBLrel. 22, Last annotation update)  
DE Hypothetical protein LAGE-2 (Fragment).  
GN Name=LAGE-2;  
OS Homo sapiens (Human).  
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
OC Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini; Hominidae;  
OC Homo.  
OX NCBI\_TaxID=9606;  
RN [1]  
RP NUCLEOTIDE SEQUENCE.  
RA Lethe B.G.;  
RL Submitted (MAR-2000) to the EMBL/GenBank/DBJ databases.  
DR EMBL; AJ275978; CAB76945.1; -; mRNA.  
KW Hypothetical protein.  
FT NON\_TER 1  
SQ SEQUENCE 142 AA; 13895 MW; 27BBE922AC4ACC7B CRC64;  
  
Query Match 60.6%; Score 581; DB 2; Length 142;  
Best Local Similarity 94.8%; Pred. No. 2.5e-34;  
Matches 109; Conservative 2; Mismatches 4; Indels 0; Gaps 0;  
  
QY 27 PGGNAGGPGGAGATGGRPGAGARASGPGGAGPRGPHGGAASGLNGCCRCGARGPESR 86  
DB 1 PGGNAGGPGGAGATGGRPGAGARASGPGGAGPRGPHGGAASGLNGCCRCGARGPESR 60  
  
QY 87 LLEFLYLPAMPATPMEALARRSLAQAADPLPVPGVLLKFTVSGNILTRLTAAD 141  
DB 61 LLEFLYLPAMPATPMEALARRSLAQAADPLPVPGVLLKFTVSGNILTMSVQDQD 115  
  
RESULT 5  
ITBA2\_HUMAN STANDARD; PRT; 143 AA.  
ID ITBA2\_HUMAN STANDARD;  
AC Q14657; Q81Z78;  
DT 01-NOV-1997 (Rel. 35, Created)  
DT 25-OCT-2004 (Rel. 45, Last sequence update)  
DT 10-MAY-2005 (Rel. 47, Last annotation update)  
DE ITBA2 protein (ESO-3 protein) (DXS9879E).  
GN Name=ITBA2; Synonyms=ESO3;  
OS Homo sapiens (Human).  
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
OC Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini; Hominidae;  
OC Homo.  
OX NCBI\_TaxID=9606;  
RN [1]  
RP NUCLEOTIDE SEQUENCE [LARGE SCALE MRNA].  
RC TISSUE=Blood, and Pancreas;  
RX MEDLINE=22388257; PubMed=12477932; DOI=10.1073/pnas.242603899;  
RA Strausberg R.D., Feingold E.A., Grouse L.H., Derge J.G., Schuler G.D.,  
RA Klausner R.D., Collins F.S., Wagner L., Shenmen C.M., Schuler G.D.,  
RA Altschul S.F., Zeeberg B., Buetow K.H., Schaefer C.F., Bhat N.K.,  
RA Hopkins R.F., Jordan H., Moore T., Max S.I., Wang J., Hsieh F.,  
RA Diachenko L., Marusina K., Farmer A.A., Rubin G.M., Hong L.,  
RA Stapleton M., Soares M.B., Bonaldo M.F., Casavant T.L., Scheetz T.E.,  
RA Brownstein M.J., Usdin T.B., Roshylyuk S., Carninci P., Frange C.,  
RA Raha S.S., Loquellano N.A., Peters G.J., Abramson R.D., Mullany S.J.,  
RA Bosak S.A., McEwan P.J., McKernan K.J., Malek J.A., Gunaratne P.H.,  
RA Richards S., Worley K.C., Hale S., Garcia A.M., Gay L.J., Hulyk S.W.,  
RA Villalón D.K., Muzny D.M., Sodergren E.J., Lu X., Gibbs R.A.,  
RA Fahy J., Helton E., Kettman M., Madan A., Rodriguez S., Sanchez A.,  
RA Blakesley R.W., Touchman J.W., Green E.D., Dickson M.C.,  
RA Rodriguez A.C., Grimwood J., Schmutz J., Myers R.M.,  
RA Butterfield Y.S.N., Krzywinski M.I., Skalska U., Smalilus D.E.,  
RA Schnerch A., Schein J.E., Jones S.J.M., Marra M.A.;  
RT "Generation and initial analysis of more than 15,000 full-length human  
RT and mouse cDNA sequences.";  
RL Proc. Natl. Acad. Sci. U.S.A. 99:16899-16903(2002).  
RN [2]



[illegible]

OC Eukaryota; Metazoa; Mollusca; Bivalvia; Pteriomorpha; Mytiloidea;  
 OC Mytiloidea; Mytilidae; Mytilinae; Mytilus.  
 OK NCBI\_TaxID=29158;  
 RN [1]  
 RP NUCLEOTIDE SEQUENCE.  
 RX MEDLINE=22038007; PubMed=12042339;  
 RA Lucas J.M., Vaccaro E., Waite J.H.;  
 RT "A molecular, morphometric and mechanical comparison of the structural  
 RT elements of byssus from Mytilus edulis and Mytilus  
 RT galloprovincialis.";  
 RL J. Exp. Biol. 205:1807-1817(2002).  
 DR EMBL; AF448526; AAM34601.1; -; mRNA.  
 DR GO; GO:0005737; C:cytoplasm; IEA.  
 DR GO; GO:0006817; P:phosphate transport; IEA.  
 DR InterPro; IPR008160; Collagen.  
 DR Pfam; PF01391; Collagen; 6.  
 KW Collagen.  
 SQ SEQUENCE 922 AA; 80750 MW; CA281AE0975BB7D4 CRC64;  
  
 Query Match 17.3%; Score 165.5; DB 2; Length 922;  
 Best Local Similarity 47.6%; Pred. No. 0.00099;  
 Matches 39; Conservative 1; Mismatches 29; Indels 13; Gaps 3;  
  
 QY 5 GRGTGSGTGDADGPGGPIPDGPGNA--GGPGAGATGCGPRGAGAAASGPGGAPR 62  
 DB 174 GGTGCG-EGQPDGPGGPGGPGGPGGPGGPGGPGGPGGPGGPGGPGGPGGPGG 232  
  
 QY 63 GPHGGAASGLNGCCRCGARGPE 84  
 DB 233 GPPGPA-----GPGGPE 244  
  
 RESULT 13  
 Q6RF45 ARATH  
 ID Q6RF45 ARATH PRELIMINARY; PRT; 420 AA.  
 AC Q6RF45;  
 DT 05-JUL-2004 (TrEMBLrel. 27, Created)  
 DT 05-JUL-2004 (TrEMBLrel. 27, Last sequence update)  
 DT 01-FEB-2005 (TrEMBLrel. 29, Last annotation update)  
 DE Hypothetical protein.  
 GN ORFNames=At2g04170;  
 OS Arabidopsis thaliana (Mouse-ear cress).  
 OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;  
 OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicotyledons;  
 OC rosids; eurosids II; Brassicales; Brassicaceae; Arabidopsi.  
 OK NCBI\_TaxID=3702;  
 RN [1]  
 RP NUCLEOTIDE SEQUENCE.  
 RA Xiao Y., Smith S.R., Ishmael N., Ayele M., Kumar N., Redman J.,  
 RA Riedmuller S., Utterback T., Whitelaw C.A., Fraser C.M., Town C.D.;  
 RL Submitted (DEC-2003) to the EMBL/GenBank/DBJ databases.  
 RN [2]  
 RP NUCLEOTIDE SEQUENCE.  
 RA Underwood B.A., Xiao Y., Moskal W., Torian U., Redman J., Wu H.C.,  
 RA Utterback T., Town C.D.;  
 RT "Arabidopsis thaliana ORF clones of hypothetical genes.";  
 RL Submitted (JUN-2004) to the EMBL/GenBank/DBJ databases.  
 DR EMBL; AY501356; AAR99368.1; -; mRNA.  
 DR EMBL; AY649317; AAT69234.1; -; mRNA.  
 DR InterPro; IPR002083; MATH.  
 DR Pfam; PF0917; MATH; 2.  
 DR SMART; SM00061; MATH; 2.  
 DR PROSITE; PS0144; MATH; 2.  
 KW Hypothetical protein.  
 SQ SEQUENCE 420 AA; 44404 MW; 6B108942B389EEA9 CRC64;  
  
 Query Match 16.9%; Score 162; DB 2; Length 420;  
 Best Local Similarity 50.0%; Pred. No. 0.00086;  
 Matches 42; Conservative 0; Mismatches 32; Indels 10; Gaps 5;  
  
 QY 5 GRGTG---GSTGDADGPGGPI--PDGPGNAGGPGAGATGCGPRGAG--AARASGPGG 59  
 DB 19 GRGGGPGGPGGPGGPGGPGGPGGPGGPGGPGGPGGPGGPGGPGGPGGPGGPGG 74



**This Page Blank (uspto)**

GenCore version 5.1.7  
Copyright (c) 1993 - 2006 Bioceleration Ltd.

OM protein - protein search, using sw model

Run on: February 4, 2006, 12:17:37 ; Search time 50 Seconds  
(without alignments)  
297.632 Million cell updates/sec

Title: US-09-529-206E-4

Perfect score: 959

Sequence: 1 MQAERGTTGGTGADGPGG.....WITQCFVFLAQPFGQRR 180

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 572060 seqs, 82675679 residues

Total number of hits satisfying chosen parameters: 572060

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%  
Maximum Match 100%  
Listing first 45 summaries

Database :

- Issued Patents AA:\*  
1: /cgn2\_6/prodata/1/iaa/5 COMB.pap.\*  
2: /cgn2\_6/prodata/1/iaa/6 COMB.pap.\*  
3: /cgn2\_6/prodata/1/iaa/H COMB.pap.\*  
4: /cgn2\_6/prodata/1/iaa/PCTUS COMB.pap.\*  
5: /cgn2\_6/prodata/1/iaa/RE COMB.pap.\*  
6: /cgn2\_6/prodata/1/iaa/backfiles1.pap.\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	959	100.0	180	1 US-08-791-495-9	Sequence 9, Appli
2	959	100.0	180	2 US-09-392-714-25	Sequence 25, Appl
3	959	100.0	180	2 US-09-165-546D-15	Sequence 15, Appl
4	959	100.0	180	2 US-09-341-829A-9	Sequence 9, Appli
5	959	100.0	180	2 US-09-849-602-30	Sequence 30, Appl
6	953	99.4	180	2 US-08-937-263B-8	Sequence 8, Appli
7	953	99.4	180	2 US-09-751-798-8	Sequence 8, Appli
8	814	84.9	180	1 US-08-791-495-7	Sequence 7, Appli
9	814	84.9	180	2 US-09-341-829A-7	Sequence 7, Appli
10	589	61.4	210	1 US-08-791-495-5	Sequence 5, Appli
11	589	61.4	210	2 US-09-341-829A-5	Sequence 5, Appli
12	160	16.7	745	1 US-09-010-928B-28	Sequence 28, Appl
13	160	16.7	870	1 US-09-010-928B-2	Sequence 2, Appli
14	156.5	16.3	633	1 US-08-642-255-73	Sequence 73, Appl
15	156.5	16.3	1065	1 US-08-642-255-72	Sequence 72, Appl
16	155.5	16.2	235	1 US-08-529-190B-1	Sequence 1, Appli
17	155.5	16.2	641	2 US-09-249-585A-3	Sequence 3, Appli
18	155.5	16.2	641	2 US-09-410-399-4	Sequence 4, Appli
19	151.5	15.8	1057	2 US-08-931-820-1	Sequence 1, Appli
20	151.5	15.8	1057	2 US-10-153-469A-16	Sequence 16, Appl
21	151.5	15.8	1057	2 US-10-153-469A-20	Sequence 20, Appl
22	151.5	15.8	1057	2 US-10-104-889-16	Sequence 16, Appl
23	151.5	15.8	1057	2 US-10-104-889-20	Sequence 20, Appl
24	151.5	15.8	1107	2 US-10-153-469A-11	Sequence 11, Appl
25	151.5	15.8	1107	2 US-10-104-889-11	Sequence 11, Appl
26	151.5	15.8	1169	2 US-10-153-469A-6	Sequence 6, Appli
27	151.5	15.8	1169	2 US-10-104-889-6	Sequence 6, Appli

28	151.5	15.8	1171	2 US-10-153-469A-8	Sequence 8, Appli
29	151.5	15.8	1171	2 US-10-104-889-8	Sequence 8, Appli
30	151.5	15.8	1341	2 US-08-963-825-18	Sequence 18, Appl
31	151.5	15.8	1341	2 US-09-500-811-18	Sequence 18, Appl
32	151.5	15.8	1341	2 US-09-570-573-18	Sequence 18, Appl
33	151.5	15.8	1341	2 US-09-548-608-18	Sequence 18, Appl
34	151.5	15.8	1388	2 US-10-153-469A-10	Sequence 10, Appl
35	151.5	15.8	1388	2 US-10-104-889-10	Sequence 10, Appl
36	151.5	15.8	1461	2 US-09-585-887-9	Sequence 9, Appli
37	151.5	15.8	1461	2 US-09-289-578-9	Sequence 9, Appli
38	151.5	15.8	1464	2 US-09-331-347C-21	Sequence 21, Appl
39	150.5	15.7	907	1 US-09-010-928B-4	Sequence 4, Appli
40	150	15.6	1008	2 US-09-219-849-8	Sequence 8, Appli
41	150	15.6	1065	1 US-08-642-255-80	Sequence 80, Appl
42	150	15.6	1065	2 US-08-642-246-16	Sequence 16, Appl
43	150	15.6	1065	2 US-09-451-208-16	Sequence 16, Appl
44	150	15.6	1065	4 PCT-US96-06229-16	Sequence 16, Appl
45	149	15.5	166	2 US-09-297-269-41	Sequence 41, Appl

ALIGNMENTS

RESULT 1  
US-08-791-495-9  
; Sequence 9, Application US/08791495  
; Patent No. 5811519  
; GENERAL INFORMATION:  
; APPLICANT: Leth, Bernard  
; APPLICANT: Lucas, Sophie  
; APPLICANT: De Smet, Charles  
; APPLICANT: Godelaine, Daniele  
; APPLICANT: Boon-Falleur, Thierry  
; TITLE OF INVENTION: LL-1 TUMOR SPECIFIC GENES  
; NUMBER OF SEQUENCES: 14  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: Wolf, Greenfield & Sacks, P.C.  
; STREET: 600 Atlantic Avenue  
; CITY: Boston  
; STATE: MA  
; COUNTRY: USA  
; ZIP: 02210  
; COMPUTER READABLE FORM:  
; MEDIUM TYPE: Floppy disk  
; COMPUTER: IBM PC compatible  
; OPERATING SYSTEM: PC-DOS/MS-DOS  
; SOFTWARE: Patent In Release #1.0, Version #1.25  
; CURRENT APPLICATION DATA:  
; APPLICATION NUMBER: US/08/791,495  
; FILING DATE:  
; CLASSIFICATION: 435  
; ATTORNEY/AGENT INFORMATION:  
; NAME: Van Amsterdam, John R.  
; REGISTRATION NUMBER: 40,212  
; REFERENCE/DOCKET NUMBER: L0461/7005  
; TELECOMMUNICATION INFORMATION:  
; TELEPHONE: 617-720-3500  
; TELEFAX: 617-720-2441  
; INFORMATION FOR SEQ ID NO: 9:  
; SEQUENCE CHARACTERISTICS:  
; LENGTH: 180 amino acids  
; TYPE: amino acid  
; TOPOLOGY: linear  
; MOLECULE TYPE: protein  
US-08-791-495-9

Query Match 100.0%; Score 959; DB 1; Length 180;  
Best Local Similarity 100.0%; Pred. No. 9.7e-79;  
Matches 180; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
QY 1 MQAERGTTGGTGADGPGGCPIDPGGNAGGPBAGATGGRGPRGAGARASGPGGA 60  
DB 1 MQAERGTTGGTGADGPGGCPIDPGGNAGGPBAGATGGRGPRGAGARASGPGGA 60

QY 61 PRGPHGAASGLNGCCRCGARGPESRLLEFYLLAMPATPMEAEARRSLAQDAPPLPVPG 120  
DE 61 PRGPHGAASGLNGCCRCGARGPESRLLEFYLLAMPATPMEAEARRSLAQDAPPLPVPG 120  
QY 121 VLLKFTVSGNLTIRLTAAADHRQLQSLSSCQLQSLMLWITQCFPLPVFLAQPPSGQRR 180  
DB 121 VLLKFTVSGNLTIRLTAAADHRQLQSLSSCQLQSLMLWITQCFPLPVFLAQPPSGQRR 180

## RESULT 2

US-09-392-714-25  
; Sequence 25, Application US/09392714A  
; Patent No. 6886147  
; GENERAL INFORMATION:  
; APPLICANT: Scanlan, Matthew J.  
; APPLICANT: Gure, Ali O.  
; APPLICANT: Williamson, Barbara  
; APPLICANT: Chen, Yao-Taeng  
; APPLICANT: Old, Lloyd J.  
; TITLE OF INVENTION: Cancer Associated Antigens and Uses  
; FILE REFERENCE: L0461/7062  
; CURRENT APPLICATION NUMBER: US/09/392,714A  
; CURRENT FILING DATE: 1999-09-09  
; EARLIER APPLICATION NUMBER: PCT/US98/14679  
; EARLIER FILING DATE: 1998-07-15  
; NUMBER OF SEQ ID NOS: 30  
; SOFTWARE: FastSeq for Windows Version 3.0  
; SEQ ID NO 25  
; LENGTH: 180  
; TYPE: PRT  
; ORGANISM: Homo sapiens  
US-09-392-714-25

Query Match 100.0%; Score 959; DB 2; Length 180;  
Best Local Similarity 100.0%; Pred. No. 9.7e-79;  
Matches 180; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
QY 1 MQAEGRTGGSTGDADGPGGPGIPDGPNGAGPGEAGATGGRPGAGAAASGPGGGA 60  
DB 1 MQAEGRTGGSTGDADGPGGPGIPDGPNGAGPGEAGATGGRPGAGAAASGPGGGA 60  
QY 61 PRGPHGAASGLNGCCRCGARGPESRLLEFYLLAMPATPMEAEARRSLAQDAPPLPVPG 120  
DB 61 PRGPHGAASGLNGCCRCGARGPESRLLEFYLLAMPATPMEAEARRSLAQDAPPLPVPG 120  
QY 121 VLLKFTVSGNLTIRLTAAADHRQLQSLSSCQLQSLMLWITQCFPLPVFLAQPPSGQRR 180  
DB 121 VLLKFTVSGNLTIRLTAAADHRQLQSLSSCQLQSLMLWITQCFPLPVFLAQPPSGQRR 180

## RESULT 3

US-09-165-546D-15  
; Sequence 15, Application US/09165546D  
; Patent No. 6723832  
; GENERAL INFORMATION:  
; APPLICANT: Knuth, Alexander; Jager, Elke; Chen, Yao, Scanlan, Matt;  
; Gure, Ali, Old, Lloyd, Ritter, Gerd  
; TITLE OF INVENTION: ISOLATED PEPTIDES CORRESPONDING TO AMINO ACID  
; SEQUENCES OF NY-ESO-1, WHICH BIND TO MHC CLASS I AND MHC CL  
; USES THEREOF  
; NUMBER OF SEQUENCES: 15  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: FULBRIGHT & JAWORSKI LLP  
; STREET: 666 Fifth Avenue  
; CITY: New York City  
; STATE: New York  
; COUNTRY: USA  
; ZIP: 10158  
; COMPUTER READABLE FORM:  
; MEDIUM TYPE: Diskette, 3.5 inch, 144 kb storage  
; COMPUTER: IBM

; OPERATING SYSTEM: PC-DOS  
; SOFTWARE: Word  
; CURRENT APPLICATION DATA:  
; APPLICATION NUMBER: US/09/165,546D  
; FILING DATE: 02-Oct-1998  
; CLASSIFICATION: 530  
; PRIOR APPLICATION DATA:  
; APPLICATION NUMBER: 09/062,422  
; FILING DATE: April 17, 1998  
; APPLICATION NUMBER: 08/937,263  
; FILING DATE: September 15, 1997  
; APPLICATION NUMBER: US 08/725,182  
; FILING DATE: October 3, 1996  
; ATTORNEY/AGENT INFORMATION:  
; NAME: Hanson, No. 6723832man D.  
; REGISTRATION NUMBER: 30,946  
; REFERENCE/DOCKET NUMBER: LUD 2166.4 CIP (09807811)  
; TELECOMMUNICATION INFORMATION:  
; TELEPHONE: (212) 318-3000  
; TELEFAX: (212) 318-3400  
; INFORMATION FOR SEQ ID NO: 15:  
; SEQUENCE CHARACTERISTICS:  
; LENGTH: 180 amino acids  
; TYPE: amino acid  
; TOPOLOGY: linear  
; SEQUENCE DESCRIPTION: SEQ ID NO: 15  
US-09-165-546D-15

Query Match 100.0%; Score 959; DB 2; Length 180;  
Best Local Similarity 100.0%; Pred. No. 9.7e-79;  
Matches 180; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
QY 1 MQAEGRTGGSTGDADGPGGPGIPDGPNGAGPGEAGATGGRPGAGAAASGPGGGA 60  
DB 1 MQAEGRTGGSTGDADGPGGPGIPDGPNGAGPGEAGATGGRPGAGAAASGPGGGA 60  
QY 61 PRGPHGAASGLNGCCRCGARGPESRLLEFYLLAMPATPMEAEARRSLAQDAPPLPVPG 120  
DB 61 PRGPHGAASGLNGCCRCGARGPESRLLEFYLLAMPATPMEAEARRSLAQDAPPLPVPG 120  
QY 121 VLLKFTVSGNLTIRLTAAADHRQLQSLSSCQLQSLMLWITQCFPLPVFLAQPPSGQRR 180  
DB 121 VLLKFTVSGNLTIRLTAAADHRQLQSLSSCQLQSLMLWITQCFPLPVFLAQPPSGQRR 180

## RESULT 4

US-09-341-829A-9  
; Sequence 9, Application US/09341829A  
; Patent No. 6794131  
; GENERAL INFORMATION:  
; APPLICANT: Leth, Bernard  
; APPLICANT: Lucas, Sophie  
; APPLICANT: De Smet, Charles  
; APPLICANT: Godelaine, Daniele  
; APPLICANT: Boon-Falleur, Thierry  
; TITLE OF INVENTION: LAGE-1 TUMOR ASSOCIATED NUCLEIC ACIDS  
; FILE REFERENCE: L0461/7066  
; CURRENT APPLICATION NUMBER: US/09/341,829A  
; CURRENT FILING DATE: 1999-10-18  
; PRIOR APPLICATION NUMBER: US 08/791,495  
; PRIOR APPLICATION NUMBER: PCT/US98/01445  
; NUMBER OF SEQ ID NOS: 14  
; SOFTWARE: FastSeq for Windows Version 3.0  
; SEQ ID NO 9  
; LENGTH: 180  
; TYPE: PRT  
; ORGANISM: Homo sapiens  
US-09-341-829A-9

Query Match 100.0%; Score 959; DB 2; Length 180;  
Best Local Similarity 100.0%; Pred. No. 9.7e-79;  
Matches 180; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy	1	MQAEGRTGSGTDADGPGGPGI	PDGPGNAGGPGCEAGATGGRGPRGAGAAARAGPGGGA	60
Db	1	MQAEGRTGSGTDADGPGGPGI	PDGPGNAGGPGCEAGATGGRGPRGAGAAARAGPGGGA	60
Qy	61	PRGPHGGAAGLNGCCRCGARGPESRLLEFYLA	MPFATPMEAEIARRSLAQDAPPLVPVG	120
Db	61	PRGPHGGAAGLNGCCRCGARGPESRLLEFYLA	MPFATPMEAEIARRSLAQDAPPLVPVG	120
Qy	121	VLLKETVSGNIIITRLTAADHROQLSIS	SSCLOQLSLMIITOCFLPVFLA	180
Db	121	VLLKETVSGNIIITRLTAADHROQLSIS	SSCLOQLSLMIITOCFLPVFLA	180

## RESULT 5

```

US-09-849-602-30
; Sequence 30, Application US/09849602
; Patent No. 6794501
; GENERAL INFORMATION:
; APPLICANT: Scanlan, Matthew J.
; APPLICANT: Old, Lloyd J.
; APPLICANT: Stockert, Elisabeth
; APPLICANT: Chen, Yao-Tseng
; TITLE OF INVENTION: Colon Cancer Antigen Panel
; FILE REFERENCE: L0461/7105(JRV)
; CURRENT APPLICATION NUMBER: US/09/849,602
; CURRENT FILING DATE: 2001-05-04
; NUMBER OF SEQ ID NOS: 30
; SOFTWARE: PatentIn version 3.0
; SEQ ID NO 30
; LENGTH: 180
; TYPE: PRT
; ORGANISM: Homo sapiens
US-09-849-602-30

```

Query Match	100.0%	Score	959	DB	2	Length	180
Best Local Similarity	100.0%	Pred. No.	9.7e-79				
Matches	180	Conservative	0	Mismatches	0	Indels	0
						Gaps	0

Qy	1	MOAEGRTGSGTGDADGCGPGI	PGPCGNAGGCEAGATCGRPGCAGHARASGPGCGA	60
Db	1	MOAEGRTGSGTGDADGCGPGI	PGPCGNAGGCEAGATCGRPGCAGHARASGPGCGA	60
Qy	61	PRGPHGGAASLNGCCRCGARGPESRLLE	FYFLAMPFATPMEAEIARLSLAQDAPPIVPVG	120
Db	61	PRGPHGGAASLNGCCRCGARGPESRLLE	FYFLAMPFATPMEAEIARLSLAQDAPPIVPVG	120
Qy	121	VLLKEFTVSGNIIITRLTAADRHQRLQLST	SSCLOQLSLMMITOCFLVFVLAQPPSQRR	180
Db	121	VLLKEFTVSGNIIITRLTAADRHQRLQLST	SSCLOQLSLMMITOCFLVFVLAQPPSQRR	180

## RESULT 6

```

US-08-937-263B-8
; Sequence 8, Application US/08937263B
; Patent No. 6274145
;
; GENERAL INFORMATION:
;
; APPLICANT: Chen, Yao-Tseung; Scanlan, Matthew;
; APPLICANT: Gure, Ali; Old, Lloyd J.; Jager, Elke;
; APPLICANT: Alexander, Knuth; Drijfhout, Jan W.
;
; TITLE OF INVENTION: ISOLATED NUCLEIC ACID MOLECULE
;
; TITLE OF INVENTION: ENCODING CANCER ASSOCIATED ANTIGEN, THE ANTIGEN
;
; TITLE OF INVENTION: ITSELF, AND USES THEREOF
;
; NUMBER OF SEQUENCES: 8
;
; CORRESPONDENCE ADDRESS:
;
; ADDRESSEE: Fulbright & Jaworski, L.L.P.
;
; STREET: 666 Fifth Avenue
;
; CITY: New York City
;
; STATE: New York
;
; COUNTRY: USA
;
; ZIP: 10103
;
; COMPUTER READABLE FORM:
;
; MEDIUM TYPE: Diskette, 3.5 inch, 144 kb storage
;
; COMPUTER: IBM

```

```

? OPERATING SYSTEM: PC-DOS
? SOFTWARE: WordPerfect
? CURRENT APPLICATION DATA:
? APPLICATION NUMBER: US/08/937,263B
? FILING DATE: September 15, 1997
? PRIOR APPLICATION DATA:
? APPLICATION NUMBER: 08/725,182
? FILING DATE: October 3, 1996
? ATTORNEY/AGENT INFORMATION:
? NAME: Sinn, Eric, Patent Agent
? REGISTRATION NUMBER: 40,177
? REFERENCE/DOCKET NUMBER: LUD 5466.1
? TELECOMMUNICATION INFORMATION:
? TELEPHONE: (212) 318-3000
? TELEFAX: (212) 752-5958
? INFORMATION FOR SEQ ID NO: 8:
? SEQUENCE CHARACTERISTICS:
? LENGTH: 180
? TYPE: amino acid
? TOPOLOGY: linear
?
? US-08-937-263B-8

```

	Query Match	99.4%	Score 953;	DB 2;	Length 180;
	Best Local Similarity	99.4%;	Pred. No. 3.4e-78;		
	Matches 179;	Conservative 0;	Mismatches 1;	Indels 0;	Gaps 0;
Qy	1	MQAEGRTGSGTGDADGPGGPGI	PDGPGGNAGGPGGEAGATGG	PRCAGAGAAASGPGCGA	60
Db	1	MQAEGRTGSGTGDADGPGGPGI	PDGPGGNAGGPGGEAGATGG	PRCAGAGAAASGPGCGA	60
Qy	61	PRGPHGGAASGLNGCCRCGARGPESRLLEFY	LAMPFATPMEAEIARRSLAQDAPPL	VPVG	120
Db	61	PRGPHGGAASGLNGCCRCGARGPESRLLEFY	LAMPFATPMEAEIARRSLAQDAPPL	VPVG	120
Qy	121	VLLKEFTVSGNIIITRLTAADHRLQQLSIS	CSCLQQLSLLMWITOCFLPVFLAQPPSGQR	180	
Db	121	VLLKEFTVSGNIIITRLTAADHRLQQLSIS	CSCLQQLSLLMWITOCFLPVFLAQPPSGQR	180	

## RESULT 7

US-09-751-798-8  
Sequence 8, Application US/09751798  
Patent No. 6525177  
GENERAL INFORMATION:  
APPLICANT: Stockert, Elisabeth; Jager, Elke;  
APPLICANT: Chen, Yao-tseng; Scanlan, Matthew;  
APPLICANT: Knuth, Alexander; Old, Lloyd J.  
TITLE OF INVENTION: Antibodies Which Bind to NY-ESO-1 Cancer  
TITLE OF INVENTION: Associated Proteins, Uses Thereof,  
TITLE OF INVENTION: Truncated Forms of NY-ESO-1, and HLA  
TITLE OF INVENTION: Binding Peptides Derived Therefrom  
NUMBER OF SEQUENCES: 8  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: Fulbright & Jaworski, L.L.P.  
STREET: 666 Fifth Avenue  
CITY: New York City  
STATE: New York  
COUNTRY: USA  
ZIP: 10103  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Diskette, 3.5 inch, 144 kb storage  
COMPUTER: IBM  
OPERATING SYSTEM: PC-DOS  
SOFTWARE: WordPerfect  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/09751,798  
FILING DATE:  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: 09/062,422  
FILING DATE:  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: 08/725,182  
FILING DATE: October 3, 1996

```
; ATTORNEY/AGENT INFORMATION:
; NAME: Hanson, No. 6525177man D.
; REGISTRATION NUMBER: 30,946
; REFERENCE/DOCKET NUMBER: LUD 5466.3
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (212) 318-3168
; TELEFAX: (212) 752-5958
; INFORMATION FOR SEQ ID NO: 8:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 180
; TYPE: amino acid
; TOPOLOGY: linear
;
US-09-751-798-8

Query Match          99.4%; Score 953; DB 2; Length 180;
Best Local Similarity 99.4%; Pred. No. 3.4e-78;
Matches 179; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 MQAEGRTGGTGDADGPGGPGIPDGGNAGGPGGAGATGGRGPRGAGAAASGPGGGA 60
   |||||
Db 1 MQAEGRTGGTGDADGPGGPGIPDGGNAGGPGGAGATGGRGPRGAGAAASGPGGGA 60

QY 61 PRGPHGGAASGLNGCCRCGARGPESRLLEFYLAMPFATPMEAEELARRSLAQDAPPLVPFG 120
   |||||
Db 61 PRGPHGGAASGLNGCCRCGARGPESRLLEFYLAMPFATPMEAEELARRSLAQDAPPLVPFG 120

QY 121 VLLKEFTVSGNLTIRLTAADHRQLQLSISSCLQLSLLMWITQCFLPVFLAQPSPGQRR 180
   |||||
Db 121 VLLKEFTVSGNLTIRLTAADHRQLQLSISSCLQLSLLMWITQCFLPVFLAQPSPGQRR 180

RESULT 8
US-08-791-495-7
; Sequence 7, Application US/08791495
; Patent No. 5811519
; GENERAL INFORMATION:
; APPLICANT: Leth, Bernard
; APPLICANT: Lucas, Sophie
; APPLICANT: De Smet, Charles
; APPLICANT: Godelaine, Daniele
; APPLICANT: Boon-Falleur, Thierry
; TITLE OF INVENTION: LL-1 TUMOR SPECIFIC GENES
; NUMBER OF SEQUENCES: 14
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Wolf, Greenfield & Sacks, P.C.
; STREET: 600 Atlantic Avenue
; CITY: Boston
; STATE: MA
; COUNTRY: USA
; ZIP: 02210
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; OPERATING SYSTEM: IBM PC compatible
; SOFTWARE: Patent In Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/791,495
; FILING DATE:
; CLASSIFICATION: 435
; ATTORNEY/AGENT INFORMATION:
; NAME: Van Amsterdam, John R.
; REGISTRATION NUMBER: 40,212
; REFERENCE/DOCKET NUMBER: L0461/7005
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 617-720-3500
; TELEFAX: 617-720-2441
; INFORMATION FOR SEQ ID NO: 7:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 180 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: protein
;
US-08-791-495-7

; ATTORNEY/AGENT INFORMATION:
; NAME: Hanson, No. 6525177man D.
; REGISTRATION NUMBER: 30,946
; REFERENCE/DOCKET NUMBER: LUD 5466.3
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (212) 318-3168
; TELEFAX: (212) 752-5958
; INFORMATION FOR SEQ ID NO: 8:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 180
; TYPE: amino acid
; TOPOLOGY: linear
;
US-09-751-798-8

Query Match          84.9%; Score 814; DB 1; Length 180;
Best Local Similarity 84.4%; Pred. No. 1e-65;
Matches 152; Conservative 13; Mismatches 15; Indels 0; Gaps 0;

QY 1 MQAEGRTGGTGDADGPGGPGIPDGGNAGGPGGAGATGGRGPRGAGAAASGPGGGA 60
   |||||
Db 1 MQAEGRTGGTGDADGPGGPGIPDGGNAGGPGGAGATGGRGPRGAGAAASGPGGGA 60

QY 61 PRGPHGGAASGLNGCCRCGARGPESRLLEFYLAMPFATPMEAEELARRSLAQDAPPLVPFG 120
   |||||
Db 61 PRGPHGGAASGLNGCCRCGARGPESRLLEFYLAMPFATPMEAEELARRSLAQDAPPLVPFG 120

QY 121 VLLKEFTVSGNLTIRLTAADHRQLQLSISSCLQLSLLMWITQCFLPVFLAQPSPGQRR 180
   |||||
Db 121 VLLKEFTVSGNLTIRLTAADHRQLQLSISSCLQLSLLMWITQCFLPVFLAQPSPGQRR 180

RESULT 9
US-09-341-829A-7
; Sequence 7, Application US/09341829A
; Patent No. 6794131
; GENERAL INFORMATION:
; APPLICANT: Leth, Bernard
; APPLICANT: Lucas, Sophie
; APPLICANT: De Smet, Charles
; APPLICANT: Godelaine, Daniele
; APPLICANT: Boon-Falleur, Thierry
; TITLE OF INVENTION: LAGE-1 TUMOR ASSOCIATED NUCLEIC ACIDS
; FILE REFERENCE: L0461/7066
; CURRENT APPLICATION NUMBER: US/09/341,829A
; CURRENT FILING DATE: 1999-10-18
; PRIOR APPLICATION NUMBER: US 08/791,495
; PRIOR APPLICATION NUMBER: PCT/US98/01445
; NUMBER OF SEQ ID NOS: 14
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 7
; LENGTH: 180
; TYPE: PRT
; ORGANISM: Homo sapiens
;
US-09-341-829A-7

Query Match          84.9%; Score 814; DB 2; Length 180;
Best Local Similarity 84.4%; Pred. No. 1e-65;
Matches 152; Conservative 13; Mismatches 15; Indels 0; Gaps 0;

QY 1 MQAEGRTGGTGDADGPGGPGIPDGGNAGGPGGAGATGGRGPRGAGAAASGPGGGA 60
   |||||
Db 1 MQAEGRTGGTGDADGPGGPGIPDGGNAGGPGGAGATGGRGPRGAGAAASGPGGGA 60

QY 61 PRGPHGGAASGLNGCCRCGARGPESRLLEFYLAMPFATPMEAEELARRSLAQDAPPLVPFG 120
   |||||
Db 61 PRGPHGGAASGLNGCCRCGARGPESRLLEFYLAMPFATPMEAEELARRSLAQDAPPLVPFG 120

QY 121 VLLKEFTVSGNLTIRLTAADHRQLQLSISSCLQLSLLMWITQCFLPVFLAQPSPGQRR 180
   |||||
Db 121 VLLKEFTVSGNLTIRLTAADHRQLQLSISSCLQLSLLMWITQCFLPVFLAQPSPGQRR 180

RESULT 10
US-08-791-495-5
; Sequence 5, Application US/08791495
; Patent No. 5811519
; GENERAL INFORMATION:
; APPLICANT: Leth, Bernard
; APPLICANT: Lucas, Sophie
; APPLICANT: De Smet, Charles
; APPLICANT: Godelaine, Daniele
; APPLICANT: Boon-Falleur, Thierry
; TITLE OF INVENTION: LL-1 TUMOR SPECIFIC GENES
; NUMBER OF SEQUENCES: 14
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Wolf, Greenfield & Sacks, P.C.
```

```
STREET: 600 Atlantic Avenue
CITY: Boston
STATE: MA
COUNTRY: USA
ZIP: 02210
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patent In Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/791,495
FILING DATE:
CLASSIFICATION: 435
ATTORNEY/AGENT INFORMATION:
NAME: Van Amsterdam, John R.
REGISTRATION NUMBER: 40,212
REFERENCE/DOCKET NUMBER: L0461/7005
TELECOMMUNICATION INFORMATION:
TELEPHONE: 617-720-3500
TELEFAX: 617-720-2441
INFORMATION FOR SEQ ID NO: 5:
SEQUENCE CHARACTERISTICS:
LENGTH: 210 amino acids
TYPE: amino acid
TOPOLOGY: linear
MOLECULE TYPE: protein
US-08-791-495-5

Query Match 61.4%; Score 589; DB 1; Length 210;
Best Local Similarity 76.6%; Pred. No. 2e-45;
Matches 108; Conservative 15; Mismatches 18; Indels 0; Gaps 0;

QY 1 MQAEGRTGGSTGDADPGGPGIPDGGNAGGPGGAGATGGRGPRGAGAAASGPGGGA 60
DB 1 MQAEGGRTGGSTGDADPGGPGIPDGGNAGGPGGAGATGGRGPRGAGAAASGPGGGA 60
QY 61 PRGPHGGAASGLNGCCRCGARGPESRLLEFYLPAMPFATPMEAEALARRSLAQAADAPLPVPG 120
DB 61 PRGPHGGAASGLNGCCRCGARGPESRLLEFYLPAMPFATPMEAEALARRSLAQAADAPLPVPG 120
QY 121 VLLKFTVSGNLTIRLTAAD 141
DB 121 AVLKDFTVSGNLLFMSVRDQD 141

RESULT 12
US-09-010-928B-28
; Sequence 28, Application US/09010928B
; Patent No. 5994099
; GENERAL INFORMATION:
; APPLICANT: Lewis, Randolph V
; APPLICANT: Havaashi, Cheryl Y
; TITLE OF INVENTION: EXTREMELY ELASTIC SPIDER SILK PROTEIN AND DNA
; TITLE OF INVENTION: CODING THEREFOR
; NUMBER OF SEQUENCES: 29
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: BIRCH, STEWART, KOLASCH & BIRCH
; STREET: 8110 GATEHOUSE RD. SUITE 500E
; CITY: FALLS CHURCH
; STATE: VIRGINIA
; COUNTRY: UNITED STATES OF AMERICA
; ZIP: 22042
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patent In Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/010,928B
; FILING DATE: 22-JAN-1998
; CLASSIFICATION: 435
; ATTORNEY/AGENT INFORMATION:
; NAME: Murphy Jr., Gerald M
; REGISTRATION NUMBER: 28977
; REFERENCE/DOCKET NUMBER: 1447-109P
; INFORMATION FOR SEQ ID NO: 28:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 745 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: protein
; US-09-010-928B-28

Query Match 16.7%; Score 160; DB 1; Length 745;
Best Local Similarity 52.4%; Pred. No. 2.9e-06;
Matches 43; Conservative 1; Mismatches 32; Indels 6; Gaps 5;

QY 5 GRGTGGSTGDADPGGPGIPDGGNAGGP-GEAGATGGRGPRGAGAAASGPGG-GAPR 62
DB 508 GPGGFGGPGGFGGPGGAGGPGYGG-AGGPGGAGGPGYGGAGG--PYGPGGAGGPGY 564
QY 63 GPHG-GAASGLNGCCRCGARGP 83
DB 565 GPGGAGGSGYLGAGGSGGVGP 586

RESULT 13
US-09-010-928B-2
; Sequence 2, Application US/09010928B
; Patent No. 5994099
; GENERAL INFORMATION:
; APPLICANT: Lewis, Randolph V
; APPLICANT: Havaashi, Cheryl Y
; TITLE OF INVENTION: EXTREMELY ELASTIC SPIDER SILK PROTEIN AND DNA
```





**This Page Blank (uspto)**

GenCore version 5.1.7  
Copyright (c) 1993 - 2006 Bioceleration Ltd.

OM protein - protein search, using sw model

Run on: February 4, 2006, 12:29:28 ; Search time 174 Seconds  
(without alignments)  
432.237 Million cell updates/sec

Title: US-09-529-206E-4

Perfect score: 959

Sequence: 1 MQAEGRGTTGGTGDADPGG.....WITQCFLPVFLAQPSPGQRR 180

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 1867569 seqs, 417829326 residues

Total number of hits satisfying chosen parameters: 1867569

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

Published Applications AA\_Main:

- 1: /cgn2\_6/prodata/1/pubpaa/US07\_PUBCOMB.pep.\*
- 2: /cgn2\_6/prodata/1/pubpaa/US08\_PUBCOMB.pep.\*
- 3: /cgn2\_6/prodata/1/pubpaa/US09\_PUBCOMB.pep.\*
- 4: /cgn2\_6/prodata/1/pubpaa/US10\_PUBCOMB.pep.\*
- 5: /cgn2\_6/prodata/1/pubpaa/US10B\_PUBCOMB.pep.\*
- 6: /cgn2\_6/prodata/1/pubpaa/US11\_PUBCOMB.pep.\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	959	100.0	180	3	US-09-849-602-30
2	959	100.0	180	4	US-10-207-655-71
3	959	100.0	180	4	US-10-026-066-3
4	959	100.0	180	4	US-10-117-937-74
5	959	100.0	180	4	US-10-295-027-386
6	959	100.0	180	4	US-10-296-734-832
7	959	100.0	180	4	US-10-188-832-139
8	959	100.0	180	4	US-10-777-053-11
9	959	100.0	180	4	US-10-751-088-15
10	959	100.0	180	4	US-10-657-022-74
11	959	100.0	180	4	US-10-837-217-11
12	959	100.0	180	5	US-10-877-373-9
13	959	100.0	180	5	US-10-723-860-1270
14	959	100.0	180	5	US-10-871-708-7
15	959	100.0	180	5	US-10-895-523-3
16	959	100.0	180	5	US-10-182-506A-3
17	959	100.0	180	5	US-10-756-149-5024
18	959	100.0	180	6	US-11-067-064-74
19	959	100.0	180	6	US-11-067-159-74
20	959	100.0	337	3	US-09-821-883-27
21	959	100.0	379	6	US-11-144-912-27
22	954	99.5	179	5	US-10-482-029-202
23	953	99.4	180	3	US-09-751-798-8
24	953	99.4	180	4	US-10-023-182-8
25	918	95.7	180	4	US-10-364-614-14
26	821	85.6	180	4	US-10-117-937-75
27	821	85.6	180	4	US-10-657-022-75

28	821	85.6	180	6	US-11-067-064-75	Sequence 75, Appl
29	821	85.6	180	6	US-11-067-159-75	Sequence 75, Appl
30	814	84.9	180	4	US-10-146-473-69	Sequence 69, Appl
31	814	84.9	180	4	US-10-296-734-834	Sequence 834, App
32	814	84.9	180	4	US-10-468-406-4	Sequence 4, Appl
33	814	84.9	180	5	US-10-877-373-7	Sequence 7, Appl
34	631.5	65.8	135	4	US-10-295-027-388	Sequence 141, App
35	631.5	65.8	135	4	US-10-188-832-141	Sequence 388, App
36	596	62.1	210	4	US-10-157-031-88	Sequence 88, Appl
37	595	62.0	210	4	US-10-117-937-76	Sequence 76, Appl
38	595	62.0	210	4	US-10-657-022-76	Sequence 76, Appl
39	595	62.0	210	6	US-11-067-159-76	Sequence 76, Appl
40	595	62.0	210	6	US-11-067-159-76	Sequence 5, Appl
41	589	61.4	210	5	US-10-877-373-5	Sequence 17, Appl
42	532	55.5	123	4	US-10-777-053-17	Sequence 17, Appl
43	532	55.5	123	4	US-10-837-217-17	Sequence 17, Appl
44	532	55.5	179	4	US-10-777-053-20	Sequence 20, Appl
45	532	55.5	179	4	US-10-837-217-20	Sequence 20, Appl

ALIGNMENTS

RESULT 1

US-09-849-602-30  
; Sequence 30, Application US/09849602  
; Publication No. US20030165834A1  
; GENERAL INFORMATION:  
; APPLICANT: Scanlan, Matthew J.  
; APPLICANT: Old, Lloyd J.  
; APPLICANT: Stockert, Elisabeth  
; APPLICANT: Chen, Yao-Tseng  
; TITLE OF INVENTION: Colon Cancer Antigen Panel  
; FILE REFERENCE: L0461/7105(JRV)  
; CURRENT APPLICATION NUMBER: US/09/849,602  
; NUMBER OF SEQ ID NOS: 30  
; NUMBER OF SEQ ID NOS: 30  
; SEQ ID NO 30  
; LENGTH: 180  
; TYPE: PRT  
; ORGANISM: Homo sapiens  
US-09-849-602-30

Query Match	100.0%	Score 959;	DB 3;	Length 180;
Best Local Similarity	100.0%	Pred. No. 1.8e-67;		
Matches 180;	Conservative 0;	Mismatches 0;	Indels 0;	Gaps 0;
QY	1	MQAEGRGTTGGTGDADPGGPGIPDPGGNAGGPGGEAGATGGRGPRGAGAAASGPGGGA	60	
Db	1	MQAEGRGTTGGTGDADPGGPGIPDPGGNAGGPGGEAGATGGRGPRGAGAAASGPGGGA	60	
QY	61	PRGPHGGAASGLNGCCRGARGPESRLLLEFYLPAMPFATPMEAEIARRSLAQADAPLPVPG	120	
Db	61	PRGPHGGAASGLNGCCRGARGPESRLLLEFYLPAMPFATPMEAEIARRSLAQADAPLPVPG	120	
QY	121	VLLKEFTVSGNLTIRLTAADHRQLQSLSSCLQLSLMLWITQCFLPVFLAQPSPGQRR	180	
Db	121	VLLKEFTVSGNLTIRLTAADHRQLQSLSSCLQLSLMLWITQCFLPVFLAQPSPGQRR	180	

RESULT 2

US-10-207-655-71  
; Sequence 71, Application US/10207655  
; Publication No. US20030118592A1  
; GENERAL INFORMATION:  
; APPLICANT: Ledbetter, Jeffrey A.  
; APPLICANT: Hayden-Ledbetter, Martha S.  
; TITLE OF INVENTION: BINDING DOMAIN-IMMUNOGLOBULIN FUSION PROTEINS  
; FILE REFERENCE: 390069.401C1  
; CURRENT APPLICATION NUMBER: US/10/207,655  
; CURRENT FILING DATE: 2002-07-25  
; NUMBER OF SEQ ID NOS: 426

```
; SOFTWARE: PatentIn version 3.0
; SEQ ID NO 71
; LENGTH: 180
; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-207-655-71

Query Match          100.0%; Score 959; DB 4; Length 180;
Best Local Similarity 100.0%; Pred. No. 1.8e-67; Indels 0; Gaps 0;
Matches 180; Conservative 0; Mismatches 0;

Qy 1 MQAEGRTGSGTGDADPGGPGIPDGGNAGGPGGAGATGGRPGAGAAASGPGGGA 60
    |||||
Db 1 MQAEGRTGSGTGDADPGGPGIPDGGNAGGPGGAGATGGRPGAGAAASGPGGGA 60
    |||||

Qy 61 PRGPHGGAASGLNGCCRCGARGPESRLLEFYLAMPPATPMEAEARSLAODAPPLVPVG 120
    |||||
Db 61 PRGPHGGAASGLNGCCRCGARGPESRLLEFYLAMPPATPMEAEARSLAODAPPLVPVG 120
    |||||

Qy 121 VLLKEFTVSGNLTIRLTAAADHRQLQLSISSCLQQLSLLMWITQCFLPVFLAQPSPGQRR 180
    |||||
Db 121 VLLKEFTVSGNLTIRLTAAADHRQLQLSISSCLQQLSLLMWITQCFLPVFLAQPSPGQRR 180
    |||||

RESULT 3
US-10-026-066-3
; Sequence 3, Application US/10026066
; Publication No. US20030215425A1
; GENERAL INFORMATION:
; APPLICANT: Simard, John J. L.
; APPLICANT: Diamond, David C.
; TITLE OF INVENTION: EPI TOPE SYNCHRONIZATION IN ANTIGEN
; FILE REFERENCE: CTLIMM.21CPLC
; CURRENT APPLICATION NUMBER: US/10/026,066
; PRIOR FILING DATE: 2001-12-07
; PRIOR APPLICATION NUMBER: 09/561,074
; PRIOR FILING DATE: 2000-04-28
; PRIOR APPLICATION NUMBER: 09/560,465
; PRIOR FILING DATE: 2000-04-28
; PRIOR APPLICATION NUMBER: 09/561,572
; PRIOR FILING DATE: 2000-04-28
; PRIOR APPLICATION NUMBER: 09/561,571
; PRIOR FILING DATE: 2000-04-28
; PRIOR APPLICATION NUMBER: PCT/US01/13806
; PRIOR FILING DATE: 2001-04-27
; NUMBER OF SEQ ID NOS: 89
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 3
; LENGTH: 180
; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-026-066-3

Query Match          100.0%; Score 959; DB 4; Length 180;
Best Local Similarity 100.0%; Pred. No. 1.8e-67; Indels 0; Gaps 0;
Matches 180; Conservative 0; Mismatches 0;

Qy 1 MQAEGRTGSGTGDADPGGPGIPDGGNAGGPGGAGATGGRPGAGAAASGPGGGA 60
    |||||
Db 1 MQAEGRTGSGTGDADPGGPGIPDGGNAGGPGGAGATGGRPGAGAAASGPGGGA 60
    |||||

Qy 61 PRGPHGGAASGLNGCCRCGARGPESRLLEFYLAMPPATPMEAEARSLAODAPPLVPVG 120
    |||||
Db 61 PRGPHGGAASGLNGCCRCGARGPESRLLEFYLAMPPATPMEAEARSLAODAPPLVPVG 120
    |||||

Qy 121 VLLKEFTVSGNLTIRLTAAADHRQLQLSISSCLQQLSLLMWITQCFLPVFLAQPSPGQRR 180
    |||||
Db 121 VLLKEFTVSGNLTIRLTAAADHRQLQLSISSCLQQLSLLMWITQCFLPVFLAQPSPGQRR 180
    |||||

RESULT 4
US-10-117-937-74
```

```
; Sequence 74, Application US/10117937
; Publication No. US20030220239A1
; GENERAL INFORMATION:
; APPLICANT: CTL IMMUNO THERAPIES CORP.
; APPLICANT: SIMARD, John, J.L.
; APPLICANT: DIAMOND, David, C.
; APPLICANT: LIU, Liping
; APPLICANT: XIE, Zhidong
; TITLE OF INVENTION: EPI TOPE SEQUENCES
; FILE REFERENCE: CTLIMM.027A
; CURRENT APPLICATION NUMBER: US/10/117,937
; PRIOR FILING DATE: 2002-04-04
; PRIOR APPLICATION NUMBER: US 60/282,211
; PRIOR FILING DATE: 2001-04-06
; PRIOR APPLICATION NUMBER: US 60/337,017
; PRIOR FILING DATE: 2001-11-07
; PRIOR APPLICATION NUMBER: US 60/363,210
; PRIOR FILING DATE: 2002-03-07
; NUMBER OF SEQ ID NOS: 602
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 74
; LENGTH: 180
; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-117-937-74

Query Match          100.0%; Score 959; DB 4; Length 180;
Best Local Similarity 100.0%; Pred. No. 1.8e-67; Indels 0; Gaps 0;
Matches 180; Conservative 0; Mismatches 0;

Qy 1 MQAEGRTGSGTGDADPGGPGIPDGGNAGGPGGAGATGGRPGAGAAASGPGGGA 60
    |||||
Db 1 MQAEGRTGSGTGDADPGGPGIPDGGNAGGPGGAGATGGRPGAGAAASGPGGGA 60
    |||||

Qy 61 PRGPHGGAASGLNGCCRCGARGPESRLLEFYLAMPPATPMEAEARSLAODAPPLVPVG 120
    |||||
Db 61 PRGPHGGAASGLNGCCRCGARGPESRLLEFYLAMPPATPMEAEARSLAODAPPLVPVG 120
    |||||

Qy 121 VLLKEFTVSGNLTIRLTAAADHRQLQLSISSCLQQLSLLMWITQCFLPVFLAQPSPGQRR 180
    |||||
Db 121 VLLKEFTVSGNLTIRLTAAADHRQLQLSISSCLQQLSLLMWITQCFLPVFLAQPSPGQRR 180
    |||||

RESULT 5
US-10-295-027-386
; Sequence 386, Application US/10295027
; Publication No. US20030232350A1
; GENERAL INFORMATION:
; APPLICANT: Afar, Daniel
; APPLICANT: Aziz, Natasha
; APPLICANT: Gineberg, Wendy M.
; APPLICANT: Gish, Kurt C.
; APPLICANT: Glynn, Richard
; APPLICANT: Hevezi, Peter A.
; APPLICANT: Mack, David H.
; APPLICANT: Murray, Richard
; APPLICANT: Watson, Susan R.
; APPLICANT: Eos Biotechnology, Inc.
; TITLE OF INVENTION: Methods of Diagnosis of Cancer, Compositions and
; FILE REFERENCE: 018501-012500US
; CURRENT APPLICATION NUMBER: US/10/295,027
; CURRENT FILING DATE: 2002-11-13
; PRIOR APPLICATION NUMBER: US 09/663,733
; PRIOR FILING DATE: 2000-09-15
; PRIOR APPLICATION NUMBER: US 60/350,666
; PRIOR FILING DATE: 2001-11-13
; PRIOR APPLICATION NUMBER: US 60/335,394
; PRIOR FILING DATE: 2001-11-15
; PRIOR APPLICATION NUMBER: US 60/332,464
; PRIOR FILING DATE: 2001-11-21
; PRIOR APPLICATION NUMBER: US 60/334,393
; PRIOR FILING DATE: 2001-11-29
```

```

Db      121  VLLKFTVSGNIIIRLTAAHROQLSISCSCLQQLSLMWITQCFLPVLQPPSGQRR 180

RESULT 7
US-10-188-832-139
; Sequence 139, Application US/10188832
; Publication No. US20040076955A1
; GENERAL INFORMATION:
; APPLICANT: Mack, David H.
; APPLICANT: Aziz, Natasha
; APPLICANT: Eos Biotechnology, Inc.
; TITLE OF INVENTION: Methods of Diagnosis of Bladder Cancer, Compositions
; TITLE OF INVENTION: and Methods of Screening for Modulators of Bladder
; TITLE OF INVENTION: Cancer
; FILE REFERENCE: 018501-002330US
; CURRENT APPLICATION NUMBER: US/10/188,832
; CURRENT FILING DATE: 2002-11-22
; PRIOR APPLICATION NUMBER: US 60/302,814
; PRIOR FILING DATE: 2001-07-03
; PRIOR APPLICATION NUMBER: US 60/310,099
; PRIOR FILING DATE: 2001-08-03
; PRIOR APPLICATION NUMBER: US 60/343,705
; PRIOR FILING DATE: 2001-11-08
; PRIOR APPLICATION NUMBER: US 60/350,666
; PRIOR FILING DATE: 2001-11-13
; PRIOR APPLICATION NUMBER: US 60/372,246
; PRIOR FILING DATE: 2002-04-12
; NUMBER OF SEQ ID NOS: 207
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 139
; LENGTH: 180
; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-188-832-139

```

```

Query Match          100.0%; Score 959; DB 4; Length 180;
Best Local Similarity 100.0%; Pred. No. 1.8e-67;
Matches 180; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY      1  MQAEGRTGSGTDADPGGPGIPDGPNGAGPGEGAGATGGRGPRGAGAARASGPGGGA 60
      |||
Db       1  MQAEGRTGSGTDADPGGPGIPDGPNGAGPGEGAGATGGRGPRGAGAARASGPGGGA 60
      |||

QY      61  PRGPHGGAASGLNGCCRCGARGPESRLLEFYLAMPPATPMEAEIARRSLAQDAPPLPVP 120
      |||
Db       61  PRGPHGGAASGLNGCCRCGARGPESRLLEFYLAMPPATPMEAEIARRSLAQDAPPLPVP 120
      |||

QY      121  VLLKFTVSGNIIITRLTAAHRQLQLSISCCQLQLSLMWITOCFLPVFLAQPPSGQR 180
      |||
Db       121  VLLKFTVSGNIIITRLTAAHRQLQLSISCCQLQLSLMWITOCFLPVFLAQPPSGQR 180
      |||

RESULT 8
US-10-777-053-11
; Sequence 11, Application US/10777053
; Publication No. US20040132088A1
; GENERAL INFORMATION:
; APPLICANT: Simard, John J. L.
; APPLICANT: Diamond, David C.
; APPLICANT: Giu, Xiyong
; APPLICANT: Lei, Xiang-Dong
; TITLE OF INVENTION: EXPRESSION VECTORS ENCODING EPITOPES OF
; TITLE OF INVENTION: TARGET-ASSOCIATED ANTIGENS AND METHODS FOR THEIR DESIGN
; FILE REFERENCE: MANNN.022C1
; CURRENT APPLICATION NUMBER: US/10/777,053
; CURRENT FILING DATE: 2004-02-10
; PRIOR APPLICATION NUMBER: 10/292,413
; PRIOR FILING DATE: 2002-11-07
; PRIOR APPLICATION NUMBER: 60/336,968
; PRIOR FILING DATE: 2001-11-07
; NUMBER OF SEQ ID NOS: 979
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 11

```

```
; LENGTH: 180
; TYPE: PRT
; ORGANISM: Homo Sapien
US-10-777-053-11

Query Match      100.0%; Score 959; DB 4; Length 180;
Best Local Similarity 100.0%; Pred. No. 1.8e-67;
Matches 180; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 MQAEGRTGGTGADGPGGPGIPDPGGNAGGPGGAGATGGRGPRGAGAAASGPGGGA 60
    |||
Db 1 MQAEGRTGGTGADGPGGPGIPDPGGNAGGPGGAGATGGRGPRGAGAAASGPGGGA 60
    |||

QY 61 PRGPHGGAASGLNGCCRCGARGPESRLLEFYLLAMPFATPMEAEALARRSLAQDAPPLVPVG 120
    |||
Db 61 PRGPHGGAASGLNGCCRCGARGPESRLLEFYLLAMPFATPMEAEALARRSLAQDAPPLVPVG 120
    |||

QY 121 VLLKEFTVSGNLTIRLTAADHRQLQLSISSCLQLSLLMWITQCFLPVFLAQPSPGQRR 180
    |||
Db 121 VLLKEFTVSGNLTIRLTAADHRQLQLSISSCLQLSLLMWITQCFLPVFLAQPSPGQRR 180
    |||

RESULT 9
US-10-751-088-15
; Sequence 15, Application US/10751088
; Publication No. US20040158044A1
; GENERAL INFORMATION:
; APPLICANT: Knuth, Alexander; Jager, Elke; Chen, Yao, Scanlan, Matt;
; Gure, Ali, Old, Lloyd, Ritter, Gerd
; TITLE OF INVENTION: ISOLATED PEPTIDES CORRESPONDING TO AMINO ACID
; SEQUENCES OF NY-ESO-1, WHICH BIND TO MHC CLASS I AND MHC CLASS II
; USES THEREOF
; NUMBER OF SEQUENCES: 15
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: FULBRIGHT & JAWORSKI LLP
; STREET: 666 Fifth Avenue
; CITY: New York City
; STATE: New York
; COUNTRY: USA
; ZIP: 10158
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Diskette, 3.5 inch, 144 kb storage
; COMPUTER: IBM
; OPERATING SYSTEM: PC-DOS
; SOFTWARE: Word
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/10751,088
; FILING DATE: 02-Jan-2004
; CLASSIFICATION: 530
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US/09/165,546D
; FILING DATE: 02-Oct-1998
; APPLICATION NUMBER: 09/062,422
; FILING DATE: April 17, 1998
; APPLICATION NUMBER: 08/937,263
; FILING DATE: September 15, 1997
; APPLICATION NUMBER: US 08/725,182
; FILING DATE: October 3, 1996
; ATTORNEY/AGENT INFORMATION:
; NAME: Hanson, Norman D.
; REGISTRATION NUMBER: 30,946
; REFERENCE/DOCKET NUMBER: LUD 2166.4 CIP (09807811)
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (212) 318-3000
; TELEFAX: (212) 318-3400
; INFORMATION FOR SEQ ID NO: 15:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 180 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; SEQUENCE DESCRIPTION: SEQ ID NO: 15
;
; JS-10-751-088-15
```

```
Query Match      100.0%; Score 959; DB 4; Length 180;
Best Local Similarity 100.0%; Pred. No. 1.8e-67;
Matches 180; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 MQAEGRTGGTGADGPGGPGIPDPGGNAGGPGGAGATGGRGPRGAGAAASGPGGGA 60
    |||
Db 1 MQAEGRTGGTGADGPGGPGIPDPGGNAGGPGGAGATGGRGPRGAGAAASGPGGGA 60
    |||

QY 61 PRGPHGGAASGLNGCCRCGARGPESRLLEFYLLAMPFATPMEAEALARRSLAQDAPPLVPVG 120
    |||
Db 61 PRGPHGGAASGLNGCCRCGARGPESRLLEFYLLAMPFATPMEAEALARRSLAQDAPPLVPVG 120
    |||

QY 121 VLLKEFTVSGNLTIRLTAADHRQLQLSISSCLQLSLLMWITQCFLPVFLAQPSPGQRR 180
    |||
Db 121 VLLKEFTVSGNLTIRLTAADHRQLQLSISSCLQLSLLMWITQCFLPVFLAQPSPGQRR 180
    |||

RESULT 10
US-10-657-022-74
; Sequence 74, Application US/10657022
; Publication No. US20040180354A1
; GENERAL INFORMATION:
; APPLICANT: Simard, John J. L.
; APPLICANT: Diamond, David C.
; APPLICANT: Liu, Liping
; APPLICANT: Liu, Zheng
; TITLE OF INVENTION: EPITOPE SEQUENCES
; FILE REFERENCE: MANNK.032A
; CURRENT APPLICATION NUMBER: US/10/657,022
; CURRENT FILING DATE: 2003-09-04
; PRIOR APPLICATION NUMBER: 60/409123
; PRIOR FILING DATE: 2002-09-06
; NUMBER OF SEQ ID NOS: 610
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 74
; LENGTH: 180
; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-657-022-74

Query Match      100.0%; Score 959; DB 4; Length 180;
Best Local Similarity 100.0%; Pred. No. 1.8e-67;
Matches 180; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 MQAEGRTGGTGADGPGGPGIPDPGGNAGGPGGAGATGGRGPRGAGAAASGPGGGA 60
    |||
Db 1 MQAEGRTGGTGADGPGGPGIPDPGGNAGGPGGAGATGGRGPRGAGAAASGPGGGA 60
    |||

QY 61 PRGPHGGAASGLNGCCRCGARGPESRLLEFYLLAMPFATPMEAEALARRSLAQDAPPLVPVG 120
    |||
Db 61 PRGPHGGAASGLNGCCRCGARGPESRLLEFYLLAMPFATPMEAEALARRSLAQDAPPLVPVG 120
    |||

QY 121 VLLKEFTVSGNLTIRLTAADHRQLQLSISSCLQLSLLMWITQCFLPVFLAQPSPGQRR 180
    |||
Db 121 VLLKEFTVSGNLTIRLTAADHRQLQLSISSCLQLSLLMWITQCFLPVFLAQPSPGQRR 180
    |||

RESULT 11
US-10-837-217-11
; Sequence 11, Application US/10837217
; Publication No. US20040203051A1
; GENERAL INFORMATION:
; APPLICANT: Simard, John J. L.
; APPLICANT: Diamond, David C.
; APPLICANT: Qiu, Zhiyong
; APPLICANT: Lei, Xiang-Dong
; TITLE OF INVENTION: EXPRESSION VECTORS ENCODING EPITOPES OF
; TARGET-ASSOCIATED ANTIGENS AND METHODS FOR THEIR DESIGN
; FILE REFERENCE: MANNK.022C2
; CURRENT APPLICATION NUMBER: US/10/837,217
; CURRENT FILING DATE: 2004-04-30
; PRIOR APPLICATION NUMBER: 10/292,413
; PRIOR FILING DATE: 2002-11-07
```

```
; PRIOR APPLICATION NUMBER: 60/336,968
; PRIOR FILING DATE: 2001-11-07
; NUMBER OF SEQ ID NOS: 979
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 11
; LENGTH: 180
; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-837-217-11

Query Match      100.0%; Score 959; DB 4; Length 180;
Best Local Similarity 100.0%; Pred. No. 1.8e-67;
Matches 180; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 MQAEGRTGSGTGDADPGCGGIPDGGNAGGPGAGATGGRPGAGAAASGPGGGA 60
Db 1 MQAEGRTGSGTGDADPGCGGIPDGGNAGGPGAGATGGRPGAGAAASGPGGGA 60
QY 61 PRGPHGGAASGLNGCCRCGARGPESRLLEFYLLAMPFATPMEALARRSLAQDAPPLVPFG 120
Db 61 PRGPHGGAASGLNGCCRCGARGPESRLLEFYLLAMPFATPMEALARRSLAQDAPPLVPFG 120
QY 121 VLLKEFTVSGNLTIRLTAADHRQLQLSISSCLQLSLMWITQCFLPVFLAQPSPGQRR 180
Db 121 VLLKEFTVSGNLTIRLTAADHRQLQLSISSCLQLSLMWITQCFLPVFLAQPSPGQRR 180

RESULT 12
US-10-877-373-9
; Sequence 9, Application US/10877373
; Publication No. US20040234541A1
; GENERAL INFORMATION:
; APPLICANT: Leth,, Bernard
; APPLICANT: Lucas, Sophie
; APPLICANT: De Smet, Charles
; APPLICANT: Godelaine, Daniele
; APPLICANT: Boon-Falleur, Thierry
; TITLE OF INVENTION: LAGE-1 TUMOR ASSOCIATED NUCLEIC ACIDS
; FILE REFERENCE: L0461/7066
; CURRENT APPLICATION NUMBER: US/10/877,373
; CURRENT FILING DATE: 2004-06-25
; PRIOR APPLICATION NUMBER: US/09/341,829
; PRIOR FILING DATE: 1999-10-18
; PRIOR APPLICATION NUMBER: US 08/791,495
; PRIOR APPLICATION NUMBER: PCT/US98/01445
; NUMBER OF SEQ ID NOS: 14
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 9
; LENGTH: 180
; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-877-373-9

Query Match      100.0%; Score 959; DB 5; Length 180;
Best Local Similarity 100.0%; Pred. No. 1.8e-67;
Matches 180; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 MQAEGRTGSGTGDADPGCGGIPDGGNAGGPGAGATGGRPGAGAAASGPGGGA 60
Db 1 MQAEGRTGSGTGDADPGCGGIPDGGNAGGPGAGATGGRPGAGAAASGPGGGA 60
QY 61 PRGPHGGAASGLNGCCRCGARGPESRLLEFYLLAMPFATPMEALARRSLAQDAPPLVPFG 120
Db 61 PRGPHGGAASGLNGCCRCGARGPESRLLEFYLLAMPFATPMEALARRSLAQDAPPLVPFG 120
QY 121 VLLKEFTVSGNLTIRLTAADHRQLQLSISSCLQLSLMWITQCFLPVFLAQPSPGQRR 180
Db 121 VLLKEFTVSGNLTIRLTAADHRQLQLSISSCLQLSLMWITQCFLPVFLAQPSPGQRR 180

RESULT 13
US-10-723-860-1270
; Sequence 1270, Application US/10723860
```

```
; Publication No. US20040253606A1
; GENERAL INFORMATION:
; APPLICANT: Aziz, Nataasha
; APPLICANT: Ginsburg, Wendy M.
; APPLICANT: Zlotnik, Albert
; TITLE OF INVENTION: Methods of Diagnosis of Soft Tissue Sarcoma, Compositions &
; FILE REFERENCE: 05882.0193.NPUS01
; CURRENT APPLICATION NUMBER: US/10/723,860
; CURRENT FILING DATE: 2003-11-26
; PRIOR APPLICATION NUMBER: 60/429,739
; PRIOR FILING DATE: 2002-11-26
; NUMBER OF SEQ ID NOS: 8393
; SOFTWARE: PatentIn version 3.2
; SEQ ID NO 1270
; LENGTH: 180
; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-723-860-1270

Query Match      100.0%; Score 959; DB 5; Length 180;
Best Local Similarity 100.0%; Pred. No. 1.8e-67;
Matches 180; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 MQAEGRTGSGTGDADPGCGGIPDGGNAGGPGAGATGGRPGAGAAASGPGGGA 60
Db 1 MQAEGRTGSGTGDADPGCGGIPDGGNAGGPGAGATGGRPGAGAAASGPGGGA 60
QY 61 PRGPHGGAASGLNGCCRCGARGPESRLLEFYLLAMPFATPMEALARRSLAQDAPPLVPFG 120
Db 61 PRGPHGGAASGLNGCCRCGARGPESRLLEFYLLAMPFATPMEALARRSLAQDAPPLVPFG 120
QY 121 VLLKEFTVSGNLTIRLTAADHRQLQLSISSCLQLSLMWITQCFLPVFLAQPSPGQRR 180
Db 121 VLLKEFTVSGNLTIRLTAADHRQLQLSISSCLQLSLMWITQCFLPVFLAQPSPGQRR 180

RESULT 14
US-10-871-708-7
; Sequence 7, Application US/10871708
; Publication No. US20050118186A1
; GENERAL INFORMATION:
; APPLICANT: Chiang, Chih-Sheng
; TITLE OF INVENTION: COMBINATIONS OF TUMOR-ASSOCIATED
; FILE REFERENCE: MANWK.035A
; CURRENT APPLICATION NUMBER: US/10/871,708
; CURRENT FILING DATE: 2004-06-17
; PRIOR APPLICATION NUMBER: 60/479,554
; PRIOR FILING DATE: 2003-06-17
; NUMBER OF SEQ ID NOS: 18
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 7
; LENGTH: 180
; TYPE: PRT
; ORGANISM: Autoimmunogenic Cancer
US-10-871-708-7

Query Match      100.0%; Score 959; DB 5; Length 180;
Best Local Similarity 100.0%; Pred. No. 1.8e-67;
Matches 180; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 MQAEGRTGSGTGDADPGCGGIPDGGNAGGPGAGATGGRPGAGAAASGPGGGA 60
Db 1 MQAEGRTGSGTGDADPGCGGIPDGGNAGGPGAGATGGRPGAGAAASGPGGGA 60
QY 61 PRGPHGGAASGLNGCCRCGARGPESRLLEFYLLAMPFATPMEALARRSLAQDAPPLVPFG 120
Db 61 PRGPHGGAASGLNGCCRCGARGPESRLLEFYLLAMPFATPMEALARRSLAQDAPPLVPFG 120
QY 121 VLLKEFTVSGNLTIRLTAADHRQLQLSISSCLQLSLMWITQCFLPVFLAQPSPGQRR 180
Db 121 VLLKEFTVSGNLTIRLTAADHRQLQLSISSCLQLSLMWITQCFLPVFLAQPSPGQRR 180
```

RESULT 15  
US-10-895-523-3  
; Sequence 3, Application US/10895523  
; Publication No. US20050130920A1  
; GENERAL INFORMATION:  
; APPLICANT: Simard, John J. L.  
; APPLICANT: Lei, Xiang-Dong  
; APPLICANT: Diamond, David C.  
; TITLE OF INVENTION: EPITOPE SYNCHRONIZATION IN ANTIGEN  
; TITLE OF INVENTION: PRESENTING CELLS  
; FILE REFERENCE: MANNK.021CF1CC1  
; CURRENT APPLICATION NUMBER: US/10/895,523  
; CURRENT FILING DATE: 2004-07-20  
; PRIOR APPLICATION NUMBER: 10/026,066  
; PRIOR FILING DATE: 2001-12-07  
; PRIOR APPLICATION NUMBER: 10/005,905  
; PRIOR FILING DATE: 2001-12-07  
; PRIOR APPLICATION NUMBER: 09/561,074  
; PRIOR FILING DATE: 2000-04-28  
; PRIOR APPLICATION NUMBER: 09/560,465  
; PRIOR FILING DATE: 2000-04-28  
; PRIOR APPLICATION NUMBER: 09/561,572  
; PRIOR FILING DATE: 2000-04-28  
; PRIOR APPLICATION NUMBER: 09/561,571  
; PRIOR FILING DATE: 2000-04-28  
; PRIOR APPLICATION NUMBER: PCT/US01/13806  
; PRIOR FILING DATE: 2001-04-27  
; PRIOR APPLICATION NUMBER: 09/999,186  
; PRIOR FILING DATE: 2001-11-07  
; NUMBER OF SEQ ID NOS: 94  
; SOFTWARE: FastSeq for Windows Version 4.0  
; SEQ ID NO 3  
; LENGTH: 180  
; TYPE: PRT  
; ORGANISM: Homo sapiens  
US-10-895-523-3

Query Match 100.0%; Score 959; DB 5; Length 180;  
Best Local Similarity 100.0%; Pred. No. 1.8e-67;  
Matches 180; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
  
Qy 1 MQAEGRTGGSTGDADGGGPGIPDGGGNAGGPGAGATGGRGPRGAGARASGPGGGA 60  
Db 1 MQAEGRTGGSTGDADGGGPGIPDGGGNAGGPGAGATGGRGPRGAGARASGPGGGA 60  
  
Qy 61 PRGPHGGAASGLNGCCRCGARGPESRLLEFYLPMPFATPMEAEIARRSLAQDAPPLPVP 120  
Db 61 PRGPHGGAASGLNGCCRCGARGPESRLLEFYLPMPFATPMEAEIARRSLAQDAPPLPVP 120  
  
Qy 121 VLLKEFTVSGNLTIRLTAAADHRQLQLSISSCLQQLSLLMWITQCFLPVLAPPSGQR 180  
Db 121 VLLKEFTVSGNLTIRLTAAADHRQLQLSISSCLQQLSLLMWITQCFLPVLAPPSGQR 180

Search completed: February 4, 2006, 12:32:52  
Job time : 175 secs

GenCore version 5.1.7  
Copyright (c) 1993 - 2006 Bioceleration Ltd.

OM protein - protein search, using sw model

Run on: February 4, 2006, 12:30:03 ; Search time 16 Seconds  
(without alignments)  
131.828 Million cell updates/sec

Title: US-09-529-206E-4

Perfect score: 959

Sequence: 1 MQAERGTCGTGADGPGG.....WITQCFPLVFLAQPSPGQRR 180

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 88029 seqs, 11718060 residues

Total number of hits satisfying chosen parameters: 88029

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Listing first 45 summaries

Database : Published Applications AA.New.\*  
1: /cgn2\_6/prodata/2/pubpaa/US08\_NEW\_PUB.pep.\*  
2: /cgn2\_6/prodata/2/pubpaa/US06\_NEW\_PUB.pep.\*  
3: /cgn2\_6/prodata/2/pubpaa/US07\_NEW\_PUB.pep.\*  
4: /cgn2\_6/prodata/2/pubpaa/PCT\_NEW\_PUB.pep.\*  
5: /cgn2\_6/prodata/2/pubpaa/US09\_NEW\_PUB.pep.\*  
6: /cgn2\_6/prodata/2/pubpaa/US10\_NEW\_PUB.pep.\*  
7: /cgn2\_6/prodata/2/pubpaa/US11\_NEW\_PUB.pep.\*  
8: /cgn2\_6/prodata/2/pubpaa/US60\_NEW\_PUB.pep.\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

#### SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	959	100.0	180	7	US-11-155-288-7
2	954	99.5	240	7	US-11-021-441-28
3	155.5	16.2	641	6	US-10-848-976-1
4	153	16.0	1079	7	US-11-052-554A-145
5	151.5	15.8	1464	7	US-11-000-463-243
6	151.5	15.8	1464	7	US-11-186-284-28
7	151.5	15.8	1464	7	US-11-021-603-2
8	151.5	15.8	1467	6	US-10-821-234-1096
9	151	15.7	1901	7	US-11-052-554A-135
10	149.5	15.6	615	7	US-11-052-554A-132
11	147	15.3	1381	7	US-11-052-554A-138
12	146.5	15.3	1767	6	US-10-995-561-911
13	146.5	15.3	1767	6	US-10-995-561-914
14	146.5	15.3	1806	6	US-10-995-561-912
15	146.5	15.3	1806	6	US-10-995-561-915
16	146.5	15.3	1818	6	US-10-995-561-910
17	146.5	15.3	1818	6	US-10-995-561-913
18	146	15.2	853	7	US-11-052-554A-149
19	145.5	15.2	1306	7	US-11-052-554A-139
20	145	15.1	741	7	US-11-052-554A-161
21	145	15.1	1329	7	US-11-052-554A-136
22	143	14.9	1660	7	US-11-052-554A-137
23	142.5	14.9	484	7	US-11-052-554A-170
24	142.5	14.9	1466	7	US-11-186-284-33
25	140.5	14.7	606	7	US-11-052-554A-163

26 138 14.4 1538 7 US-11-052-554A-146 Sequence 146, App  
27 137.5 14.3 801 7 US-11-052-554A-166 Sequence 166, App  
28 137.5 14.3 1166 6 US-10-821-234-964 Sequence 964, App  
29 135.5 14.1 749 7 US-11-052-554A-148 Sequence 148, App  
30 135.5 14.1 1366 6 US-10-821-234-1431 Sequence 1431, App  
31 135.5 14.1 1366 7 US-11-186-284-31 Sequence 31, App1  
32 134.5 14.0 914 7 US-11-052-554A-160 Sequence 160, App  
33 134 14.0 80 7 US-11-029-003-34 Sequence 34, App1  
34 133.5 13.9 3063 7 US-11-186-284-26 Sequence 26, App1  
35 133 13.9 767 7 US-11-052-554A-154 Sequence 154, App  
36 131.5 13.7 1736 7 US-11-124-368A-329 Sequence 329, App  
37 131 13.7 200 6 US-10-841-129-7 Sequence 7, App1  
38 131 13.7 200 7 US-11-060-659-32 Sequence 32, App1  
39 130.5 13.6 434 7 US-11-052-554A-167 Sequence 167, App  
40 130.5 13.6 576 7 US-11-052-554A-164 Sequence 164, App  
41 130.5 13.6 923 7 US-11-052-554A-147 Sequence 147, App  
42 129 13.5 639 7 US-11-052-554A-165 Sequence 165, App  
43 129 13.5 1496 7 US-11-186-284-35 Sequence 35, App1  
44 128.5 13.4 778 7 US-11-052-554A-144 Sequence 144, App  
45 128 13.3 618 7 US-11-052-554A-150 Sequence 150, App

#### ALIGNMENTS

RESULT 1  
US-11-155-288-7  
; Sequence 7, Application US/11155288  
; Publication No. US20060008468A1  
; GENERAL INFORMATION:  
; APPLICANT: Chiang, Chih-Sheng  
; APPLICANT: Simard, John J. L.  
; TITLE OF INVENTION: COMBINATIONS OF TUMOR-ASSOCIATED  
; TITLE OF INVENTION: ANTIGENS IN DIAGNOSTICS FOR VARIOUS TYPES OF CANCERS  
; FILE REFERENCE: MANWK.050A  
; CURRENT APPLICATION NUMBER: US/11/155,288  
; CURRENT FILING DATE: 2005-06-17  
; PRIOR APPLICATION NUMBER: 60/580,969  
; PRIOR FILING DATE: 2004-06-17  
; NUMBER OF SEQ ID NOS: 40  
; SOFTWARE: FastSeq for Windows Version 4.0  
; SEQ ID NO 7  
; LENGTH: 180  
; TYPE: PRT  
; ORGANISM: Homo sapiens  
US-11-155-288-7

Query Match 100.0%; Score 959; DB 7; Length 180;  
Best Local Similarity 100.0%; Pred. No. 2.3e-78;  
Matches 180; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
Qy 1 MQAERGTCGTGADGPGGCGIPDGGNAGCGENAGTGGPRGAGARASGPGGA 60  
Db 1 MQAERGTCGTGADGPGGCGIPDGGNAGCGENAGTGGPRGAGARASGPGGA 60  
Qy 61 PRGPHGAASGLNGCCRCGARGPESRLLEFYLLAMPATPMEALARRSLAQADAPLPVPG 120  
Db 61 PRGPHGAASGLNGCCRCGARGPESRLLEFYLLAMPATPMEALARRSLAQADAPLPVPG 120  
Qy 121 VLLKEFTVSGNLTIRLTAAHHRQLSLSSCLQLSLMWITQCFLPVFLAQPSPGQRR 180  
Db 121 VLLKEFTVSGNLTIRLTAAHHRQLSLSSCLQLSLMWITQCFLPVFLAQPSPGQRR 180  
RESULT 2  
US-11-021-441-28  
; Sequence 28, Application US/11021441  
; Publication No. US20050249748A1  
; GENERAL INFORMATION:  
; APPLICANT: DUBENSKIY, Thomas W., Jr.  
; APPLICANT: FORTNOY, Daniel A.  
; APPLICANT: LUCKETT, William S., Jr.  
; APPLICANT: COOK, David N.







Search completed: February 4, 2006, 12:33:14  
Job time : 17 secs

**This Page Blank (uspto)**